Between meaning and essence explaining necessary truth

Schieder-Hestermann, Jakob

Awarding institution:
King's College London

Between Meaning and Essence - Explaining Necessary Truth

Dissertation
zur Erlangung des akademischen Grades

Doctor philosophiae
(Dr. phil.)

eingereicht an der Philosophischen Fakultät I
der Humboldt Universität zu Berlin
im Rahmen des Joint PhD Programms mit dem King's College London

Dissertation submitted to
Philosophischen Fakultät I of Humboldt University Berlin
as part of the Joint-PhD Program between Humboldt University Berlin
and King's College London

von/*by* Jakob Schieder-Hestermann

Präsidentin der Humboldt Universität zu Berlin:
Prof. Dr. ing. Dr. Sabine Kunst

Dekanin der Philosophischen Fakultät I:
Prof. Dr. Gabriele Metzler

Gutachterinnen und Gutachter/*Examiners*:

1. Prof. Dr. Tobias Rosefeldt
2. Prof. Dr. Keith Hossack
3. Prof. Dr. Barbara Vetter

Datum der Disputation: 30. Mai 2017

Abstract

An explanation of why some truths are necessarily true needs to make intelligible how it is that a truth is guaranteed to be true. The thesis argues that a promising starting point for an explanation of necessity can be found in Kant's containment-account of analyticity, for it explains how the truth of a judgment is guaranteed by its structure and the relationship between its constituents. This, however, can merely be a starting point for a general explanation of necessary truth, for it is both too narrow, and presupposes a contentious view of concepts. The thesis thus explores how the general strategy, explaining necessary truth by certain relationships between the constituents of representations, can be expanded to cover further necessary truths, for example a posteriori necessities and essentialist claims. It is argued that the explanation can be generalized by focussing on what it is that constitutes reference between representations and the objects they represent as well as how these representations come together to form truth-evaluable representations. Necessary truth, on this account, is a property of truth-evaluable representations which a representation has in virtue of the appropriate relationship between what is required for its truth, and the way in which the reference of its constituents is determined. The appropriate relationship guarantees the truth of the representation. After applying the theory to a range of examples, interesting parallels to essentialist accounts of necessity emerge and it is argued that the proposed explanation gets the relationship between essence and necessity right and may even be used to elucidate what essences are.

Zusammenfassung

Eine Erklärung, warum manche Wahrheiten notwendigerweise wahr sind, sollte verständlich machen, warum es für diese eine Wahrheitsgarantie gibt. Einen interessanten Ansatz für eine solche Erklärung liefert Kants Definition analytischer Wahrheiten als solche, die bereits in einem Begriff enthalten sind. Die notwendige Wahrheit analytischer Sätze kann hier über das Verhältnis ihrer Bestandteile erklärt werden. Diese Erklärung kann jedoch nur der Anfang einer generellen Erklärung von Notwendigkeit sein, denn einerseits sind nicht nur analytische Urteile notwendig, sondern auch andere, andererseits basiert die Erklärung auf einer umstrittenen Theorie von Begriffen. Die Dissertation untersucht nun, ob und wie die Strategie, Notwendigkeit über das Verhältnis zwischen den Bestandteilen von Repräsentationen zu erklären dennoch ausgeweitet und verteidigt werden kann, um zum Beispiel auch a posteriori Notwendigkeiten und essentialistische Urteile zu erfassen. Indem auf die die Referenz-Relation konstituierenden Fakten Bezug genommen wird, wird gezeigt, dass eine solche Erklärung tatsächlich möglich ist. Notwendige Wahrheit ist demnach eine Eigenschaft von wahrheitsfähigen Repräsentationen, die diese aufgrund des Verhältnisses zwischen den für die Wahrheit der Repräsentation nötigen Fakten und den Fakten, die die Referenz der Bestandteile der Repräsentation bestimmen, hat. Stehen diese in einem bestimmten Verhältnis, wird die Wahrheit der Repräsentation garantiert. Nachdem dieser Ansatz auf eine Reihe von Beispielen angewendet wird, zeigen sich interessante Parallelen zu essentialistischen Theorien von Notwendigkeit und es lässt sich zeigen, dass der Ansatz das Verhältnis zwischen Essenzen und Notwendigkeit richtig darstellt und es sogar erlaubt, ein besseres Verständnis davon zu entwickeln, was Essenzen sind.

Acknowledgements

While writing a dissertation involves much time spent alone writing, it is not something that is done alone. Ideas develop in an environment of support, discussion, and friendship.
I want to thank all those people, who have made the three years of working on my PhD a wonderful intellectual and personal experience, and who contributed to the dissertation by discussing ideas, supporting me, and occasionally providing the necessary distraction.
First of all, thanks to my supervisors Barbara Vetter and Mark Textor for your continuous support, your patience with my ideas, and your constant productive criticism that has strengthened many arguments, and clarified many thoughts.
Second, I want to thank all the people who I had the opportunity to present my work to, and who helped clarify my thesis with their questions, critical remarks, and discussions after talks. These were, among others, the people at the research seminar at King's, the great people at Barbara Vetter's Colloquium in Berlin, and the audiences at various conferences, most notably: New Directions in the Metaphysics of Modality in Berlin, GAP.9 in Osnabrück, and Issues on the (Im)possible IV in Bratislava.
From the many people I met at during the PhD program, I want to thank two in particular, who have become friends and who have made my stays in London very enjoyable: Dave Jenkins and Jørgen Dyrstad.
The dissertation would not have been possible without the Joint PhD program between HU-Berlin and King's College London, which just started at the beginning of my PhD and allowed me to

spend two half-years in London. This wonderful opportunity greatly contributed to the PhD-experience.

For the last three years I received generous funding: For the first three months of the program the Carl und Max Schneider-Stiftung zur Förderung der Philosophie provided me with a stipend, before the Konrad-Adenauer-Stiftung took over for the remaining three years of the program. Thank you for the generous funding, but also for the support in the groups of stipendiaries, and the seminars on various non-philosophy-topics I could attend. There I got to meet great people from other disciplines, who have become friends.

Without my friends outside of philosophy, I would not have had the necessary distractions and determination to finish this project. So thanks to my friends from the gliding club, especially Marc, who have allowed me to sometimes just fly away. And thanks to our friends in Heidelberg and Berlin, who provided the necessary context for the thesis.

The most important people are mentioned last, my wife and my family. They were and are always there for me, and support me greatly. So thank you Johanna for the wonderful time we had, and the time we will have together, and for coming to London with me. And thanks to my family, my brother, my mom, and my dad, who I could and can call anytime, and who opened their homes whenever we would want to come by.

Contents

Introduction

The necessary is the purview of philosophy. But what is it to be necessary as opposed to being merely contingent? How is it that some things just cannot fail to be the case? The present thesis seeks to give answers to these questions.

I am not the first to attempt an answer. The last century has seen many attempts, including a rejection of the very question, ranging from a denial of the reality of necessity to the contention that necessity is fundamental and not to be further explained. Along the way preconceptions about the strict separation of the *a posteriori* natural sciences and allegedly *a priori* philosophy were swept away. In spite of much progress in the study of modality generally, a better understanding of the relationships between different modal concepts, a rich philosophical discussion, and diversity of opinions on the source of necessity, a satisfactory answer to the question what necessity is, is still lacking. I shall argue that this situation can be remedied. An illuminating, non-circular, intelligible, and general explanation of metaphysical necessity can be given.

Here is the view in a nutshell: Necessary truth is a property of truth-evaluable representations which it has in virtue of the appropriate relationship between what is required for its truth, and the way in which the reference of its constituents is determined. The appropriate relationship guarantees the truth of the representation.

This proposal, I shall argue, takes what is right about approaches to necessity which proceeded from analyticity, truth in virtue of meaning, without having to commit to conventional truth, and it takes what is right about accounts that proceed from truth in

virtue of essence, without having to commit to a primitive notion of truth in virtue of essence. Thus the theory sits comfortably between two views of necessity that seem diametrically opposed, but should better be conceived of as two faces of the same coin.

The first of these positions is an explanation of necessary truth in terms of analyticity. On this view, which was popular in the first half of the twentieth century among empiricists, necessary truth is to be equated with truth in virtue of meaning, where truth in virtue of meaning just amounts to conventional truth as opposed to factual truth.[1] This position seemed to have the benefit of accounting for necessity in a way that is compatible with empiricist principles, for coming to know about a necessary truth means knowing a linguistic convention, a knowledge that is unproblematic for members of the relevant linguistic community.

The main trouble with this view is that it misrepresents the role conventions play in contributing to truth. For all that linguistic convention can plausibly do, is to say that a sentence or word is to mean something or other, not also make it the case that what it says really is the case.

It may be that some conventions do legislate that a certain sentence is to always express a truth. Such a sentence may properly be said to be true by convention. This truth by convention, however, is not to be equated with necessary truth, for a sentence, even if it is to be used to only express truths, may express a different, contingent truth on each occasion of use. Consider "I am here now": According to the standard (and simplified) semantic treatment this sentence, whenever uttered by someone, expresses a truth. But the truth that is expressed at each

[1] For a short and clear statement of this view see Ayer (1936a).

occasion is always a different, contingent truth. This suggests that conventional truth doesn't give us what we wanted an explanation of. We wanted to know what it takes for the content of a sentence to be guaranteed to be the case, but instead we only got an explanation of how a sentence can be guaranteed to express some content that is true. So truth in virtue of meaning, even if it may, in a limited number of cases, tell us something interesting about sentences, cannot tell us anything interesting about necessary truth.

If analyticity is this easily rejected as an explanans of necessity, the natural move is to turn away form the word to the world, to the essence of things. Turning Quine's warning on its head that "meaning is what essence becomes when it is divorced from the object of reference and wedded to the word"[2] one may seek to move the explanation of necessity back to the things the necessary truths are about, or more precisely, to their essence. Necessary truths, on this view, are those which are true in virtue of the essence of objects.

The trouble with this view, however, is to make sense of the notion of truth in virtue of essence. For it seems that the same insight that already spoke against truth in virtue of meaning also speaks against truth in virtue of essence: Necessary truths, just like contingent truths, are true in virtue of what they say really being the case. It is just that what they say could not have failed to be the case. So how does the fact that some essential properties play a role in the truthmaking make it the case that a proposition is necessarily true? It is not clear how an answer may go, and so

[2] Quine 1951a, p. 22.

this problem at the very least shows that the essentialist position is in need of clarification.

The explanation of necessity to be defended in the following occupies a middle ground between these two positions: The essentialist is right that necessary truth is not just a matter of convention. Truth must be grounded in the world. Necessity is not subject to conventions in the way advertised by the conventionalist. The essentialist, however, moves the source of necessity completely away from what is said to be necessary, making it a mystery how necessary truth arises.

The analyticity-theorist, on the other hand, was right that necessary truths, in some way, do seem to be a matter of definition and meaning, but misidentifies the way meaning plays a role in giving rise to necessary truth.

The right position to take, I suggest, is the following: Necessity is not to be found exclusively in the meaning of words, nor is it to be found exclusively in the essence of things, it is to be found in the way world and representation interact.

The first step towards such a theory of necessity is to take a closer look at what the objects are which are said to be necessarily true. The analyticity theorist takes sentences to be appropriate objects, for they are subject to the conventions which allegedly guarantee their truth. However, we have seen that this focus on sentences is wrong for the purpose at hand, because conventional truth is not identical to necessary truth: a sentence that is guaranteed to express something true, is not thereby guaranteed to say something that is necessarily true. Thus, instead of taking sentences to be the objects that are necessarily true, I propose to take representations as the objects of study.

Representations, in contrast to sentences, have their representational properties essentially. They are not, however, what has classically been known as propositions, for representation tokens are not abstract objects, and they do not play the same theoretical role in a theory of meaning. Rather, a token sentence cum meaning, or a mental state with its content are examples of representation. Representations are individuated by their representational properties. Different token representations are of the same type in virtue of having the same representational properties.

Taking representations to be the objects whose necessary truth is to be explained, leads us away from a narrow picture of representation based on a communication model of language and a contingent association of words and sentences with meanings, and thus allows for a more fruitful way of considering truth and reference.

Once we help ourselves to representations, we can ask what the representational properties are whose investigation may tell us something interesting about necessary truth. The first property is, not surprisingly, truth. What is it that determines the truth of a representation? To answer this, we have to distinguish between truth-evaluable representations and referential representations. Referential representations are the constituents of truth-evaluable representations and may be either singular, referring rigidly to one object, or general, potentially referring to a number of objects. A truth-evaluable representation is true, if the referents of its constituents stand in the appropriate relation to each other. The representation expressed by the sentence “Ravens are black” for example, is true, if the referents of the referential representation expressed by ‘ravens’, are among the referents of the referential

representation expressed by ‘black’. Call the relationship between the referents of the constituent referential representations required for the truth of a truth-evaluable representation its *requirements for truth*. This is the first ingredient to the explanation of necessary truth.

The second ingredient is reference. Reference is a relation that holds between a referential representation and the object or objects it refers to. It is not a primitive relation, however. There is an explanation for why it holds. It holds in virtue of properties of the representation, call these the representation’s *reference determining properties*, and in virtue of properties of the referent itself, call these the *reference realizing properties*. What the reference realizing properties of an object are, is determined by the reference determining properties of the representation, in the sense that once it is settled what the reference determining properties of the representation are, it is also settled what properties an object would have to have to be the referent of the representation.

With requirements for truth and reference realizing properties in hand, we have everything we need for an explanation of why some representations cannot fail to be true: It so happens that sometimes the reference realizing properties determined by the constituent referential representations are related such that the requirements for truth of a truth-evaluable representation are guaranteed to be satisfied. In this case the representation cannot fail to be true, it is necessarily true.

This explanation positions itself between the essentialist and the analyticity theorist, for it recognizes that it is something about the objects that partly explains necessary truth, that is, its reference realizing properties. However, the reason these properties of the

object are so special lies in their being constitutive of reference, and so their connection to necessary truth can be explained.

The suggested explanation of necessary truth takes its inspiration from a Kantian account of analyticity, which says that a judgment consisting of a subject concept and a predicate concept is analytic, if the predicate concept is contained in the subject concept.[3] From this account an explanation of necessary truth similar to the present explanation can be given: Concept containment guarantees that the referents of the subject concept are among the referents of the predicate concept. Since this is just what is required for the truth of a subject-predicate judgment, the judgment cannot fail to be true.

The Kantian account of analyticity, however, is much too narrow to serve as a model for a general explanation of necessary truth, for concept containment is just not a very ubiquitous phenomenon. So to gain a proper understanding of necessary truth generally, a framework is needed that goes beyond simple subject-predicate judgments and concepts that contain each other. It is the task of the representation-framework to fill this role, and so to go beyond the simple case of concept containment.

So far this is a very rough sketch, but a first idea of how an explanation of necessary truth that strikes a balance between truth in virtue of essence and truth in virtue of meaning may have emerged. It is the purpose of the thesis to convince you that this explanation is a good and illuminating explanation of necessary truth. I want to use the rest of the introduction to give an overview over how the argument will proceed and give a short abstract of the points I discuss in each chapter.

---

[3] Kant's famous containment-definition of analyticity can be found in his Critique of Pure Reason (KrV A 6-7; Kant (1998)).

The thesis is divided into two parts. The first is an introductory part in which I clarify what I take to be a good philosophical explanation of necessary truth, consider the merits and challenges of previous theories of necessity, and discuss in some more detail earlier attempts at utilizing analyticity in an explanation of necessary truth.

The second part starts by laying out the broadly semantic framework in which the theory is to be spelled out. It proceeds by stating the official explanation of necessary truth in terms of the developed concepts, applies it to a number of examples, considers how the account relates to other kinds of necessity, shows how it solves a range of apparent puzzles, compares the explanation to essentialist theories of necessity, and finally argues that all this gives us reason to believe that the present explanation is a good philosophical explanation of necessary truth.

Part I

Chapter 1: Explaining Necessity

I argue that it is the task of a philosophical explanation to give an account of what it takes, constitutively, for something to be, or to be of a certain kind. So an explanation of necessary truth should tell us what it takes, constitutively, for a representation to be necessarily true. I argue that a constitutive philosophical explanation is not achieved by giving a conceptual or metaphysical reduction, even though both methods are important philosophical tools that may help in the construction of a constitutive explanation. The quality of a given philosophical explanation can be judged by considering how well the theory fares with respect to three general desiderata: non-circularity, generality, and intelligibility.

Chapter 2: Explanations of Necessity

The most prominent theories (or classes of theories) of necessity are evaluated with respect to their explanatory merits. Genuine modal realism, reductive ersatzism, primitivism, and essentialism, with the exception of reductive ersatzism, succeed in their attempt to give an explanatorily non-circular account of necessity. As such these theories do not face simple knock-down arguments. However, each theory's claim to being the best theory of necessary truth is hampered by different deficits in satisfying the intelligibility requirement. Genuine modal realism asks us, against common sense, to believe in many other concrete worlds. Different versions of primitivism, as well as essentialism, use explanantia that themselves seem in need of further explanation, so these theories only move us marginally closer to a satisfactory understanding of what it takes for a representation to be necessarily true. If a better theory of necessity is to be given, it must be non-circular, general, and score higher in terms of intelligibility than its competitors.

Chapter 3: Analyticity and Necessity

The notion of analyticity and its relation to necessity is explored. First, the main traditional problems of an account of necessary truth in terms of analyticity are laid out and suggested fixes in the literature discussed. The theories as well as the fixes are found to be lacking, since they do not solve the basic problem of the theory: Analyticity on these orthodox accounts is conceived of as a property of sentences, but the property of a sentence to be guaranteed to express a true proposition and the property of being necessarily true are different properties that apply to different

objects, and so truth in virtue of meaning cannot explain necessary truth.

More recent accounts of analyticity have therefore given up either their explanatory ambitions of explaining necessary truth, or they take a more anti-realist approach to talk of necessity. While both reactions are understandable in light of the problems faced by accounts of analyticity, looking back at the Kantian account of analyticity in terms of concept containment helps to elucidate what is right about an explanation of necessary truth in terms of analyticity. Since Kant does not conceive of analyticity as a property of sentences, but of judgments, an explanation of necessary truth suggests itself. This explanation is not nearly general enough to account for all necessary truths, but it provides a blueprint of how a more general explanation may proceed.

Part II

Chapter 4: Representation, Reference, and Truth

A semantic framework in which representations are the primary objects of study is developed. The more traditional communication model of meaning and semantics is contrasted with the new framework in which representations are considered as having their representational properties essentially. We can distinguish between truth-evaluable representations and referential representations, where referential representations are the constituents of truth-evaluable representations. Their central representational features are truth and reference. Neither truth nor reference are taken as primitive, however.

A reference relation holds in virtue of properties of the representations and in virtue of properties of the referent. The former are the *reference determining properties* and the latter are

the *reference realizing properties*. Reference realizing properties are determined by the reference determining properties. This gives us a criterion for type-identity of referential representations: Referential representations are of the same type, iff they determine the same reference realizing properties.[4]

A truth-evaluable representation is true, if the referents of the constituent referential representations are appropriately related. The relation between the referents required for the truth are the *requirements for truth* of the representation. What the respective requirements for truth are, depends on how the truth-evaluable representation is constituted by the referential representations. With the concept of requirements for truth in hand, we can give a criterion for type identity of truth-evaluable representations: Truth-evaluable representations are of the same type, iff they have constituents of the same type and the requirements for truth are the same. Finally some reasons for adopting this representation-framework for an explanation of necessity are considered.

Chapter 5: An Explanation of Necessity

The explanation of necessary truth is spelled out in detail: A truth-evaluable representation is necessarily true (if it is), because its requirements for truth are guaranteed to be satisfied by the relationship between the reference realizing properties determined by the constituent referential representations. This explanation is defended against a possible circularity worry. I argue that the proposed explanation, *prima facie*, satisfies the desiderata for a

[4] Talk of referential representations determining reference realizing properties should be taken as shorthand for reference determining properties of referential representations determining reference realizing properties.

good philosophical explanation. That it really satisfies them is argued for in the following chapters.

Chapter 6: Applying the Theory

The theory is put to use in a number of example cases. The simplest examples are, unsurprisingly, conceptual truths along the lines of Kant's own example of unmarried bachelors. The proposed explanation, however, has the resources for going beyond such simple examples and also explains the necessity of representations involving relations, as well as logical truths.

A complication and a familiar puzzle arise in cases of necessarily true representations that involve singular representations, for it seems that the necessary truth of such representations implies the necessary existence of the object the truth is about. In the present framework this puzzle and related puzzles involving mere possibilia and impossible objects, can be satisfactorily solved by clarifying what the requirements for truth of the necessarily true representations are.

Once these puzzles are solved, I argue that the theory also explains the necessary truth of essence attributions and, more tentatively, of mathematic truths.

Chapter 7: Varieties of Necessity

The proposed explanation of necessary truth is primarily an explanation of metaphysical necessity. I defend that the necessity explained is indeed metaphysical necessity, by arguing that it is the strongest kind of necessary truth which still yields an intelligible notion of possibility. I go on to propose that weaker kinds of necessary truth, for example nomic or practical necessity, can be accounted for by the familiar strategy of relativizing to a

set of background-conditions. Recent challenges to this monist account of necessity are discussed, but found wanting.

Chapter 8: Essence and Representation

The present analytic theory of necessity occupies a middle ground between theories of necessity in terms of analyticity and essentialist theories. In this chapter the challenge the essentialist faces in accounting for its central explanatory notion, truth in virtue of essence, are discussed in detail. The challenge is not easily surmountable, but the present explanation of necessity also provides an explanation of essence, thus suggesting that the explanatory relations run neither from essence to modality, nor from modality to essence, but that both are explained by a third: the elements of the present theory.

Essences, in the representation-framework, can be identified with the reference realizing properties determined by singular representations. By doing so, the role essences play in an explanation of necessary truth becomes intelligible and, in addition, an explanation of what makes some properties of objects so special as to count as essential to the object can be given. While the present theory may diverge in spirit form some essentialist intuitions, it is a means of better understanding what essences are.

Chapter 9: Evaluating the Explanation of Necessity

The final chapter summarizes the diverse points that speak in favor of the theory. I argue that there is good reason to take the theory to be true, because it provides a non-circular, illuminating and intelligible as well as general explanation of metaphysical necessity that appears extensionally adequate, accounts for

puzzles about possibilia and necessary existence, explains what essences are, and along the way makes only minimal assumptions about the nature of representation and its connection to the world. All this, I argue, makes the suggested approach well suited to further our understanding of what it takes for a truth to be necessarily true.

Part I

Chapter 1: Explaining Necessity

Philosophers typically try to explain very general phenomena: Epistemology, for example, is concerned with knowledge and justification generally, ethics is concerned with the good and right life generally, metaphysics with the general structure of the universe, and so on. A typical kind of question to ask for a philosopher in all these subfields is "what is X?" where X is some central object of inquiry in the domain in question. In epistemology a question is "what is knowledge?". In ethics it is "what is the good?". And in a branch of metaphysics it is "what is necessity?". While questions of this kind are routinely asked and answered by philosophers, it is not always explicit what would constitute a good answer to them. The diversity of answers indicates that, apart from substantial disagreement, there is no uniform understanding of the philosophical project.

It is the purpose of this first chapter to survey some of the methods one may use to answer the philosophical question "what is necessity?" and to develop some very general criteria a good answer to this question should satisfy.[5] To sharpen these general remarks, I will in the second chapter see how popular theories of modality have fared with respect to the desiderata.

1. Methods of Philosophical Investigation

Some strategies for answering questions of the form "what is X?" can be discerned. There is, for example, classical conceptual analysis, further, there are a number of similar, but at least

---

[5] These desiderata are not intended to apply to all projects taking place under the broad roof of philosophical inquiry. But they apply to the explanatory project I embark on here: explaining necessity.

apparently distinct methods: explication and metaphysical reduction. Finally, there is what we may call constitutive explanation. In the following I will introduce each of these methods, unpack some of the assumptions and requirements for such an explanation to work, argue that they can all be conceived of as being in the business of moving us towards constitutive explanations and give some general criteria of adequacy for these explanations.

Conceptual Analysis

One historically quite popular answer to philosophical questions is to give a conceptual analysis of the concept standing for some X.[6] What exactly a conceptual analysis is, is not quite clear, but, as the name says, we may think of it as proceeding from the dual assumption that the concept of X is complex, that is, that it can be analysed in the literal sense, broken down into constituents (which are themselves concepts) and that this breaking down of the concept tells us something interesting about the object of inquiry, for example knowledge, the good, necessity, etc.[7]

The first assumption is a condition for the possibility of an analysis, for if the concept to be explained is not complex, there is nothing into which it can be analyzed.[8] The second assumption is needed, because otherwise it would be unclear how conceptual analysis can tell us anything about the subject matter we are

---

[6] No assumptions about what concepts are, will currently be made. Also no complications in terms of indexical words, associated with different concepts from context to context are taken into account, as this would be an unnecessary complication in relation to the modest aim of this chapter.

[7] For an overview over different conceptions of analysis in the history of philosophy up to today see Beaney (2014).

[8] In some philosophical debates it has been claimed that no analysis is possible. Famously, Moore claimed that ‘good’ was indefinable (Moore 1903)

trying to gain insight into, for philosophy generally is interested in concepts only in virtue of the concept telling us something about the subject matter.

To get a better grip on what is involved in a conceptual analysis, let me introduce an obvious, but non-philosophical example: 'red ball'. The concept 'red ball' is quite straightforwardly a complex concept, so the first condition is satisfied. It is built up from the simpler concepts 'red' and 'ball'.[9] How does analyzing this complex concept into its constituents help us to find out something about the subject matter, that is, red balls? It tells us something about what red balls are: We come to know by the decomposition of the concept that every red ball is both a ball and that it is red. The way in which the decomposition of a concept can tell us something (necessary) about red balls has its roots in how the reference of a complex concept is determined from the reference of its constituents: The reference of 'red ball' is determined in such a way that it picks out those objects which satisfy both the condition of being red and of being a ball. So the extension of 'red ball' will be the intersection of the simpler concepts 'red' and 'ball'. We may state this in the canonical form as a (necessitated) biconditional: Something is a red ball, iff it is red and it is a ball. It seems that the present conceptual analysis has given us some answer to the admittedly not very substantial question "what is a red ball?".

While this provides a nice model for answering questions of the form "what is X?", the assumptions that have to be made point to

---

[9] There is no need at present to go into detail about how concepts come to build up each other, as this will depend on the favorite theory of concepts. With respect to the second assumption at work in conceptual analysis, we can say, however, that being built up by concepts entails that the reference of complex concepts is determined by the reference of simpler constituent concepts.

some potential problems with the universal applicability of this strategy. The most important problem is this: Are the concepts denoting the philosophically interesting phenomena complex in the way that 'red ball' is complex? It seems that there is no guarantee for this to be the case. Worse, it seems that virtually no analysis of an interesting philosophical concept has been satisfactory and so there is, *prima facie,* good reason to believe that philosophical concepts do not exhibit the structure required for conceptual analysis.[10]

Against this skeptical worry it may be objected that for reference to be possible, concepts *must* be complex and we just haven't been able to arrive at satisfactory decompositions yet. But this cannot be a good line of argument, as at least the simplest concepts from which others are built up must have their reference determined in a different way and if we need a different account of reference for them, it is not obvious that it cannot be used for more general philosophical concepts as well.

Suffice it to say, while conceptual analysis, in its pure form introduced above, seems to have an eminent tradition in philosophy, it is doubtful that it is the best and only method to arrive at answers to philosophical questions.

Before moving on to other strategies for answering philosophical questions, it is helpful to get clear on the explanatory relations at play in a conceptual analysis. I have talked of philosophical explanations, but this may be considered to be slightly odd given that the question we started out with was not a question asking why something is the case, but rather a question about what something is. But from the answer to the what question, we get

---

[10] Williamson (2002) forcefully makes this point for the concept knowledge.

answers to a number of why questions. We do, for example, get an answer to the why question why this object (pointing to a red ball) is a red ball. Answer: because it is red and a ball. If we can give an analysis of a complex concept, we come to know what it takes to be of the kind the concept describes. What it takes to be a red ball, for example, is being red and being a ball: The constitution relations holding between the complex concept and the simple concept guarantee that objects in the extension of 'red ball' are all and only the objects in the intersection of the concepts 'red' and 'ball'.

If the preconditions of a conceptual analysis are satisfied, it provides us with an answer to philosophical questions. It is doubtful, however, whether these preconditions are satisfied in the case of many philosophical concepts and so different ways of explaining what it takes for something to be a necessary truth need to be explored.

Explication

Our ordinary concepts, while some of them may be complex, are not always philosophically interesting: they are often idiosyncratic and may be ill suited for, say, scientific explanations. Consequently, analyzing them, even if possible, is not always the most interesting thing a philosopher can do. Rather, philosophers can offer helpful constructions, or reconstructions of interesting concepts in terms of well understood constituents. This is what we may, following Carnap, call the explication of a concept: "The task of making more exact a vague or not quite exact concept used in everyday life […], or rather of replacing it by a newly constructed, more exact concept,

belongs among the most important tasks of logical analysis and logical construction. We call this the task of explicating, […].”[11]

Explication, in a sense, turns conceptual analysis on its head. Instead of breaking down concepts, concepts are built up. This method seems especially adequate when considering technical concepts, which may be introduced in this way.

The answers to philosophical questions of the form “what is X?” we gain from this exercise, however, is the same as in the case of conceptual analysis and the explanatory relations run just as described above. The difference between conceptual analysis and explication lies in the way we arrive at these judgments: Instead of breaking down concepts, we construct concepts.

Since the methodology is different from conceptual analysis, some of the challenges to conceptual analysis just do not arise. There are different challenges, however, which mirror the challenges to conceptual analysis, but reflect the difference in methodology. A problem that does not arise for explications is the question whether the target concept is complex. Since the concept is explicitly constructed, it is bound to be complex. Also, there is no need for the constructed concept to match any of our pre-theoretical concepts, rather, what is required is that the concept plays the role it is supposed to play, for example in scientific explanations.

Two other problems, however, arise when explication is used in philosophical explanations. First, it is unclear how interesting a constructed concept is in each case and whether it covers the interesting cases we intended to fall into the extension of the philosophical concept in the first place. Second, the choice of well

[11] Carnap (1947), p. 7/8.

understood primitives may be too limited to construct concepts for everything we want to have a concept of, such that the method cannot be universally applied.

The reconstructive project of conceptual analysis and the constructive project of explication are rarely ever performed in their pure form. Most conceptual analyses have a revisionary, or constructive element and most explications orient themselves towards existing concepts. So both conceptual analysis and explication can be understood as complementary methods in the philosopher's toolbox. As both will yield the same kind of answer to the philosophical question of what something is, it is to be expected that both methods will serve the purpose of giving an adequate answer to philosophical questions, if the preconditions for their applicability are satisfied.

A note on biconditionals

Explications as well as conceptual analyses are usually expressed in the analytic philosopher's favorite form: as (necessitated) biconditionals. Biconditionals are a convenient way of expressing conceptual analyses as well as explications, but one needs to be careful not to take every true and necessary biconditional to express a conceptual analysis. For biconditionals by themselves, even if necessary, do not tell us anything about the explanatory relations that hold between its two sides.

To have an example where a biconditional may express a conceptual analysis, take again the example of the 'red ball' above. Since the concept 'red ball' is built up of the concepts 'red' and 'ball', the following biconditional is guaranteed to hold: Something is a red ball, iff it is red and a ball. The necessary truth of the biconditional itself does not tell us why it holds of

necessity, so it does not tell us anything about the relation that holds between the concept on the left hand side and the concepts on the right hand side of the conditional. The necessity of the conditional may be guaranteed in another way than by concept-containment. This comes out clearest in cases of necessarily coextensive concepts like ‘being the number 2’ and ‘being the smallest prime number’. Plausibly, no conceptual containment relations hold between these two concepts, but nonetheless the biconditional “something is the number 2 if and only if it is the smallest prime number” is necessarily true.

This example is likely to not be alone. Even in cases where a concept seems to be amenable to analysis, as for example in the case of ‘bachelor’ being analysed in terms of the concepts ‘unmarried’ and ‘male’, it needs to be shown that the necessary truth of the biconditional “someone is a bachelor, iff he is unmarried and male” is indeed due to conceptual containment relations that determine the reference of the complex concept in terms of the simpler ones. This is not obvious and needs to be supplemented with some conception of how concepts do come together to form complex concepts.

This illustrates that giving a philosophically interesting answer to questions of the kind “what is X?”, does not only involve finding a different way of expressing what has already been said with a biconditional, but rather to find a reformulation that allows for an adequate constitutive explanation of the phenomenon in question. What constitutive explanations are, shall be discussed in the following, but first let us look at metaphysical reduction.

Metaphysical reduction

Conceptual analyses as well as explications offer what is known as reductions: That the red ball is a red ball requires of it nothing but being red and being a ball. This reduction is achieved in virtue of the complex concept having the simpler concepts as constituents.

It is important to be clear, however, about what is reduced in a conceptual analysis: It is not the complex concept that is reduced, for it is the constitution of the concept that explains what something is. Rather it is the features of objects that fall in the extension of the concept which are reduced to other, more primitive, features. In the case of 'red ball' what is reduced is the feature of being a red ball. It is reduced to the two simpler features of being red and being a ball.

Despite this metaphysical character of the reduction flowing from a conceptual analysis, philosophers have made a distinction between conceptual reduction and metaphysical reduction, intended to extend reductions to cases in which there is no apparent conceptual connection.[12] The feature meant to be preserved in metaphysical reductions is that what is to be reduced is, in some sense, nothing over and above what it is reduced to. What is to be avoided in contrast to standard conceptual analysis, is the need for a conceptual containment relation for this reduction to be effected.

The need for metaphysical reduction comes up most urgently, if conceptual analysis is conceived of as purely *a priori*. For it seems that some true reduction-claims are not knowable *a priori*. The simplest such cases are *a posteriori* identity claims like

---

[12] See for example Hale (2013), Chapter 3, for a distinction between conceptual and metaphysical reduction, as applied to an explanation of necessary truth.

"Hesperus is Phosphorus". Hesperus is nothing over and above Phosphorous (and vice versa), after all, there is just one object. Similar examples can be given in the case of properties. Consider, for example, the putatively necessary truth that everything that is water is H_2O. Since there is apparently no conceptual containment involved in these cases, the reductive claim cannot be justified with respect to any conceptual containment relations.

However, it should be noted that at the level of metaphysics, the reductive claim in all cases, including conceptual analysis cases, is, if complete, an identity claim: being red and being a ball just is being a red ball, being H_2O just is being water, being Hesperus is being Phosphorus. The difference between conceptual analysis and metaphysical analysis lies in the way in which these identity claims are justified. In the case of conceptual analysis and explication it is concept constitution that does the trick. In the case of metaphysical reduction it is some other reason we have for believing that both concepts (water and H_2O, Hesperus and Phosphorus) denote the same objects. In the case of Hesperus and Phosphorus this may be justified by a causal theory of reference along Kripke's lines, in the case of water and H_2O some version of Putnam's theory of reference determination for natural kinds may be used to justify the identity claim.[13]

In what follows, I will, when discussing attempts at reductive theories of necessity, not distinguish between conceptual and metaphysical reductions, as there is no difference in kind of the reduction, there just is a difference in the underlying explanation of how the reference of the concepts is determined and resulting from that, a difference in their epistemological properties.

[13] see Kripke (1980) and Putnam (1975) for sketches of externalist theories of reference.

Nonetheless, what runs under the label of metaphysical reduction is a welcome addition to the philosopher's toolkit, as it widens the scope of conceptual interrelations beyond that of (arguably rare) containment-relations and moves the focus away from a mere *a priori* methodology.

2. Explanation and Reduction

Answers to philosophical questions should not only tell us how we may reformulate something with the help of some necessary biconditional, rather, they should be explanatory of what it takes to be X, where X is some philosophically interesting property like knowledge, necessity, or truth.

A philosophical explanation of this kind should not tell us how something came to be, that is, it should not be a genealogical explanation, rather, it should tell us how the thing in question is what it is constitutively.

The difference between a genealogical explanation and a constitutive explanation is best brought out by example: draw a triangle on a sheet of paper. Then ask yourself: why is there a triangle? One answer is that there is a triangle, because you just drew the triangle. This is the genealogical explanation. But there is another possible answer: there is a triangle, because there is a closed figure with straight sides and three corners on the sheet of paper. This is the constitutive explanation of why there is a triangle. What we are looking for in philosophy is constitutive explanation.

What I call constitutive explanation has in recent years gained a lot of attention in the philosophical debate under the label of

"grounding".[14] I will deliberately leave the distinction between genealogical and constitutive explanation at the intuitive level and will not enter into the quite extensive debate on grounding. The sense of explanation at issue, I take it, is clear enough by example and will be further clarified in what follows.

Given that philosophical explanation is constitutive explanation, how are conceptual analysis and metaphysical reduction adequate tools in arriving at such an explanation? In answering this question a puzzle arises right at the outset: what conceptual analyses as well as metaphysical reductions supply us with are, at bottom, identities; they give us a further way of literally saying the same thing. It is even a criterion of adequacy of reductions and analyses that what is said on either side of the biconditional is the same. But if so, how is explanation by conceptual analysis or metaphysical reduction ever possible? For explanation is commonly, and rightly, taken to be irreflexive. That is, nothing can explain itself.[15]

This problem is reminiscent of the paradox of analysis, just at the metaphysical level. The paradox of analysis goes like this: Given that it is a criterion of adequacy for philosophical analyses that the analysans has the same meaning as the analysandum, a correct analysis just amounts to an uninformative triviality, provided that meaning is known by all competent speakers of the language or users of the conceptual framework. For in this case, the identity should be transparent to all parties involved in the philosophical

---

[14] For a collection of recent work on grounding, see Correia and Schnieder (2012). Further introductory discussion can be found in Rosen (2010) and Trogdon (2013). For a more skeptical approach see Wilson (2014).

[15] The literature on grounding recognizes that problem as well and gives an answer similar to the one given here. See for example the introduction to Correia and Schnieder (2012), or see Rosen (2010).

discussion and no question about the correctness of the analysis should arise.[16] So either the analysis is trivial, or it is inadequate. Both horns of the dilemma seem unattractive.

One standard solution to the apparent paradox is to deny that giving an identity is always uninformative. For identities can be informative, if they are not knowable *a priori*, as the classic example of "Hesperus is Phosphorus" shows. So given that sameness of meaning is not always *a priori* and transparent to all involved in the debate, it may be that an adequate analysis is informative after all.

This solution to the paradox of analysis solves the problem as it was stated above. However, by just focussing on the epistemic dimension of the paradox of analysis, the proposed answer misses the somewhat deeper metaphysical dimension of the paradox. If philosophical explanation is to explain what something is, what necessity is, what knowledge is, what the good is, it does not suffice to find a different way of saying that something is necessary, or known, or good. For this would not tell us anything about the philosophical phenomenon in question, it would just tell us something about the representations we use.[17] So 'solving' the paradox by making sameness of meaning *a posteriori* is merely an epistemic fix, just as the difference between metaphysical reduction and conceptual reduction is merely an epistemic difference: We need a more thoroughly metaphysical answer to the paradox of analysis.

---

[16] A formulation of the paradox of analysis can be found in the collection "The Philosophy of G.E. Moore" in the essays by Langford (1942) and Moore (1942). For a summary see Beaney (2014).

[17] This is not to say that conceptual interconnections and identities are not philosophically interesting, it is just to say that they usually don't provide us with an answer to constitutive philosophical questions.

To see how a solution may go, let us look at an example. Assume we would come to know by conceptual analysis that knowledge just is justified, true belief. Then we may ask of a particular piece of knowledge: why is this knowledge? The answer may seem to be: because it is a justified, true belief. The answer seems adequate, because what it is to be knowledge just is being a justified, true belief. The paradox of analysis arises, because, according to the conceptual analysis, being a justified, true belief just is the same as being knowledge. So "x is a justified, true belief" and "x is knowledge" both represent the same fact. And so it may seem that when we say "something is knowledge, because it is a justified, true belief", we are saying nothing more than "something is knowledge, because it is knowledge". The irreflexivity requirement on explanation is violated.

The solution lies in taking the explanatory relation not as holding between the fact that x is knowledge and the conjunctive fact that x is a justified, true belief, for these are indeed identical according to the conceptual analysis. Rather, the explanatory relation holds between x's being knowledge and the facts that x is justified and that x is a true belief. So the explanation-relation does not connect two single identical facts, but a fact and the distinct facts it takes for it to obtain. This is even more explicit if we ask what it takes for something to be a justified, true belief. The answer is that it takes something to be *both* justified and a true belief. The conjunctive fact is explained by each of its conjuncts obtaining, not by the conjunctive fact itself.

This helps to make explicit the role of conceptual analysis and metaphysical reductions in a philosophical enquiry: It is to make explicit what it takes for something to be the case, by giving us a representation of an identical fact for which we have a clearer

idea of what it takes for it to obtain. The analysis makes transparent what it takes for something to be the case.

So conceptual analysis as well as metaphysical reduction are methods in our toolbox for arriving at constitutive explanations. However, it should be clear that they are far from being the only methods for arriving at such explanations. For it may be that either the conceptual or linguistic resources are not enough to rephrase every philosophically interesting concept, or that, even if some rephrasing is possible, it does not move us closer to seeing what it would take for something to be of the philosophically interesting kind. So when evaluating reductions, we do not only have to check whether the analysans means the same as the analysandum, we also have to check whether the analysis moves us closer to an understanding of what it takes for something to be of the philosophically interesting kind.

The understanding of philosophical explanation as giving constitutive explanations will guide the following inquiries, as it allows for an inclusive methodology and at the same time it allows for the formulation of some standards against which different proposals for explanation may be evaluated. In what follows some general criteria for evaluating an explanation of necessity will be given, before the next chapter surveys some proposals for the explanation of necessary truth.

3. The Explanation of Necessity

We can use the remarks on analysis and explanation to develop general criteria for what an explanation of necessity should accomplish. Three general desiderata for an adequate explanation of necessity can be distinguished: Non-circularity, generality, and intelligibility, broadly construed. I will discuss each in turn.

Non-circularity, as was pointed out above, is a standard requirement on explanation. It says that the explanans should not be contained in the explanandum. The special, but also most obvious case of circularity is triviality where something is directly explained by itself. Since irreflexivity is a standard requirement on explanations and non-circularity is a corollary of it, given transitivity, it should not be a controversial requirement, but also one that is not too hard to avoid.[18] The kind of circularity at issue in explanations, however, should be distinguished from circularity as it occurs in conceptual analyses. For circularity in conceptual analyses is problematic, not because it violates irreflexivity, but because no complex concept can contain itself as constituent.[19]

The second requirement is generality. Unlike the non-circularity requirement, generality is not a binary matter, but a matter of degree. It says that a philosophical account of necessity should equip us with a template for explaining why each instance of a necessary truth is a case of necessary truth. A perfectly general account of some philosophical phenomenon gives an explanation of why each fact of the relevant kind holds. A less than general account gives an explanation of only some instances of the phenomenon.

An example for a less than general explanation would be an explanation of what knowledge is that only explains knowledge gained via the auditory senses. It would thus fall short of what a

---

[18] Transitivity may not hold generally for all kinds of explanations, but arguably it holds for full philosophical explanations. For a proposal on the logic of the connective 'because' see Schnieder (2011), but also Fine (2012).

[19] Work on circularity has helped to elucidate the problematic and unproblematic features of circularity. See Keefe (2002) for some ways in which circularity may be tolerated. When applying the criterion, I will take these insights into account.

philosophical explanation of knowledge is supposed to do: giving an account of knowledge generally.

A lack of generality, however, may not always be a bad thing. It may also be indicative of a false expectation that the phenomenon in question is uniform. In such a case, the discovery of a non-general explanation may help us to understand that a phenomenon we conceive of as uniform is in fact quite disparate. Consequently, care must be taken when applying the criterion of generality in the particular case to take into account the possible disparity of the phenomenon. Still, a comparative judgment between theories is, other things equal, possible: A theory that gives a wholly general explanation is, other things equal, better than a theory which only gives a partial explanation.

The final requirement is intelligibility.[20] This is a rather loose requirement and a catch all phrase for a range of different sub-requirements, among them the classical theoretical virtues of simplicity and elegance, but also something we may call believability, the requirement that the theory should be somewhat credible. Nonetheless, it is important that these virtues are not assessed singly, but under the heading of the goal these requirements aim at: a theory that helps us to further our understanding of the world and thus makes intelligible what it takes for a certain phenomenon of philosophical interest to obtain. One important indicator for intelligibility is whether the theory requires us to take up a number of beliefs we would not have thought to be independently plausible or whether it requires us to

[20] The phrase intelligibility is used for lack of a better catch all phrase for a number of theoretical virtues. It should not be thought to imply that a theory which scores low on this requirement can literally not be understood, rather, a theory that scores high on this requirement, fits better with our overall world-view.

let go of beliefs we regard to be obvious or well justified. The reason this requirement matters is that it helps us assess how well the theory fits into our overall picture of the world and relates to our other beliefs. If it is radically revisionary in a number of ways, there is need for very strong independent reasons to believe in the theory.

The need for a good fit with beliefs is especially urgent in the area the theory is about. If a theory of knowledge, for example, classifies most of what we take to be knowledge as non-knowledge, it is probably not a good theory of knowledge, but possibly a theory of something else. Equally, if a theory of necessity classifies a lot of truths which we would take to be necessarily true as non-necessary (or vice versa) there is at least defeasible reason to believe that something about the theory is wrong.

These three criteria, non-circularity, generality, and intelligibility will be the guide in assessing different prominent theories of modality, as well as my proposal in the second part. In the following chapter the most influential contenders for a theory of modality are presented and assessed against the developed requirements.

Chapter 2: Explanations of Necessity

Some methods used for giving philosophical explanations as well as very general criteria for a successful theory of necessity were discussed. In this chapter, the abstract discussion will be made more concrete by looking at some popular approaches to the explanation of necessity. This will help to better understand the different projects that have been pursued when explaining necessity as well as their strengths and weaknesses. The discussion, however, is neither intended as a complete survey of theoretical options, nor as a refutation of any of them. The intention is to get a clearer picture of the challenges popular approaches face and what it is that may make them unsatisfactory.
I will start out with one of the most influential theories, genuine modal realism as spelled out by David Lewis, move on to (non-primitivist) Ersatz-Theories, before considering primitivism as a possible way to respond to the challenges these theories face. Finally essentialism, which has regained popularity only relatively recently, will be discussed.
Before starting, a caveat is in order. I will not discuss each position in sufficient detail to reflect all intricacies of the vast research literature on each that has been accumulating during the last half century. Rather, I will attempt to give a generic outline of each theory that gives the basic idea behind each a reading that is as charitable as possible. The lack in detail will be weighed up by the benefits for exposing the explanatory relations each theory posits and the resulting challenges they face.

1. Possible Worlds - Genuine Modal Realism

The starting point for possible world theories of modality are two biconditionals:[21]

(Nec) Necessarily p, iff in every possible world p

(Pos) Possibly p, iff in some possible world p

The two biconditionals associate talk of necessity with quantification over possible worlds, or more intuitively, over possibilities. This association has some intuitive basis in our everyday ways of speaking. We can, for example, either say that it possibly rains in the afternoon, or we can say that there is a possibility that it will rain in the afternoon. Both ways of expressing possibility seem to get the same information across. Similarly in the case of necessity: There is no possibility in which p is false, expresses the same as necessarily p.

Apart from this intuitive evidence possible worlds are the standard way of conceiving of the structures used to give formal semantics for modal expressions.[22] So it seems that there is good reason to take the biconditional above to be true.

While the truth of the biconditionals is quite uncontroversial, it is an open question, whether the right hand side of the biconditional gets us any closer to an understanding of what it takes for a proposition to be necessarily or possibly true. That is, it is not so obvious that the proposed direction of explanation is helpful as a constitutive explanation of possibility and necessity.

The proponent of genuine modal realism thinks that a genuine explanation of modality is provided by possible worlds. He believes that we should take the right hand side of the

---

[21] The theory is most famously defended by David Lewis in his (1986) and already suggested in Lewis (1973).

[22] See for example Kripke (1959) and Kripke (1963), but also Marcus (1995)

biconditional as literally quantifying over possible worlds. Truth at those possible worlds is to be understood in the same sense as truth in the actual world, just that the world in question is a different world. According to David Lewis, we should think of those other worlds as concrete individuals, just like our world, which are different worlds from the actual world in virtue of not being spatiotemporally connected to our universe.

If we intend to take talk of truth at possible worlds seriously, as explaining what it takes to be necessarily true, this surely is the way to go. For it seems to deliver the most straightforward understanding of what truth at a possible world consists in: ordinary truth at a different point in modal space.[23]

If we take the possible worlds to explain modality, however, we face a dilemma. Either we have to accept that there are a lot, really a lot, of concrete universes spatiotemporally disconnected from ours, one for every possible difference to our world, or we have to take all our possibility talk to be false and every truth as necessarily true, because we cannot bring ourselves to believe that there are all these other concrete universes.[24]

For the analysis to provide us with a good philosophical explanation, all the possible worlds really must exist and be

---

[23] To account for some intricacies involving identity across possible worlds, Lewis (1983, 1986) introduces counterpart-theory, which can be understood to be his official theory of necessity. This complication will not concern us presently, as it also implies the necessity of the biconditional above and it also requires the existence of concrete possible worlds.

[24] To get some feeling for how many concrete possible worlds there would have to be for our possibility talk to be true, think about the world in which you now have one hair less than you actually do, or a world in which there is one less grain of salt in your meal, and so on. We get a vast number of worlds, even before we get to the more ordinary possibilities, like the possible rain in the afternoon.

delimited by what is possible. But what reason do we have to believe that there are all these concrete possible worlds?

David Lewis answer to that question is the following: "Because the hypothesis is serviceable, and that is a reason to think that it is true."[25] The hypothesis is serviceable, according to Lewis, because it allows us to give a host of philosophical explanations: With possible worlds we can explain necessity and possibility. With possible worlds we can explain essences and other modal notions. With possible worlds we can explain counterfactual conditionals in terms of similarity of possible worlds. With possible worlds we can explain what propositions are. With possible worlds we can explain what properties are.[26]

Still, as Lewis admits, "Modal realism does disagree, to an extreme extent, with firm common sense opinion about what there is."[27] He takes this to be a price worth paying.

The Lewisian evaluation of the tradeoff between firm common sense opinion and alleged theoretical benefits has not been widely shared. Most philosophers have taken common sense opinion to win out and opted for a different view of what possible worlds are. The case against concrete possible worlds has further been bolstered by doubts that a possible worlds framework can deliver the benefits advertised. There is, for example, good reason to doubt that sets of possible worlds are propositions, because all necessarily true propositions would come out as identical.[28] Also,

---

[25] Lewis (1986) p.3

[26] Lewis (1986) shows how this may be done in great detail.

[27] Lewis (1986) p. 133

[28] This is a fairly standard objection to an analysis of propositions in terms of sets of possible worlds and there are some ways to alleviate the strength of the objection, but it remains the case that sets of worlds will not be fine-grained enough to serve as propositions. For an overview and a recent discussion see Hanks (2009)

essences may not be analyzable in terms of possible worlds.[29] Similar doubts may apply to the other alleged benefits of accepting possible worlds.

The details, however, will not be of further concern here, let me instead go over the desiderata for a good philosophical explanation, to see how the theory fares with respect to them.

The first desideratum for a good philosophical explanation is non-circularity. This desideratum is satisfied, even though it has been challenged by some.[30] To see this, consider how the explanation of the necessity of an arbitrary necessary truth p is achieved: We need two components for the explanation. First, we need to be told what it takes for a proposition to be necessary. This part of the explanation is given by the biconditional (nec): To be necessarily true is to be true in all possible worlds. Second, we need it to be the case that p really is true in each possible world. So given that p is necessarily true, it is necessarily true that p, because p is true in each concrete world. So necessary truth is explained by truth at multiple places in modal space. What it takes to be necessarily true is truth at all such "locations".

Some may be inclined to object, that to delimit the possible worlds, we need to make reference to what is possible, but this is a misunderstanding of Lewis' position. Which worlds exist does not depend at all on what is necessary and what is possible. To get all the possibilities we want, we have to believe that there are a lot of them, but that itself is not part of the explanation and Lewis has

---

[29] Fine (1994).

[30] For some attempts along these lines see Divers and Melia (2002) and (2003), (2006). These attempts, however, should better not be read as alleging circularity, rather as doubting that there is adequate reason to believe that there are enough possible worlds, safe by giving a modal account of them. A rejoinder can be found in Cameron (2010a) and (2012).

a way of distinguishing between different worlds that is non-modal: Different worlds are spatiotemporally disconnected from each other. So I conclude that the requirement of non-circularity does not threaten genuine modal realism in its explanatory ambitions.

The second requirement is generality, and at least if the hypothesis that there are a lot of other possible worlds is correct, it will hold. For if there are all these worlds, then the suggested explanation works quite generally for any necessary truth. It even offers a nice way to conceive of other, more restricted kinds of possibility, by quantifying over a subset of all possible worlds.[31] So there is no obvious quarrel with this requirement either.

The theory does have a fairly big deficit in satisfying the third desideratum, intelligibility, however. For genuine modal realism asks us to start believing in the existence of a lot of other concrete possible worlds. This is a fairly heavy burden, for while we might sometimes be inclined to believe in physical theories of parallel universes, we surely have no good independent reason to believe that there is a world for every way our world could be. This speaks strongly against the proposed explanation, for if we take seriously our belief that there are no other worlds apart from ours and we also take the analysis to be correct, we will arrive at the conclusion that every proposition is necessary and nothing that is not actually true is possible. This, however, seems to be plainly false, for some things that are not actually the case are surely possible.

While Lewis takes this to be a reason to believe that there are really quite a lot of different possible worlds, it is more likely to

[31] How this may be done is described by Lewis here (1986)

be taken as a reason to think that the biconditional, if it requires quantification over concrete worlds on the right hand side, is false. From this point of view, the postulated existence of a pluriverse of possible worlds seems like a classic reductio ad absurdum.[32]

A charitable way of interpreting the findings here is this: If there should be no good way to explain necessary truth without recourse to concrete possible worlds, then this constitutes some reason to believe that there are all these concrete possible worlds. However, we have no good reason to take the theory to give us a good explanation, given that we don't believe that there are all these possible worlds.

So in the following I will take Lewis by his word when he writes: "I acknowledge that my denial of common sense opinion is severe, and I think it is entirely right and proper to count that as a serious cost. […] *I* still think the price is right, high as it is. Modal realism ought to be accepted as true. The theoretical benefits are worth it. Provided, of course, that they cannot be had for less."[33] The discussion in the following is intended to show that this proviso does not hold. There is a way to have an explanation of necessity for less, and that is a reason to think that Lewis' analysis is false.

2. Possible Worlds - Non-Primitivist Ersatzism

Lewis' genuine modal realism has been influential in the discussion on modality, but it has not found many adherents. The

---

[32] Lewis (1986) acknowledges this criticism and discusses it under the label of the "incredulous stare", but ultimately dismisses it, by challenging his opponents to come up with a better theory. A detailed critique along similar lines can be found in Jubien (2008).

[33] Lewis (1986) p.135

majority of philosophers have preferred to keep the biconditionals (nec) and (pos), but to read 'possible world' as referring to some more uncontroversial representational entities, or collections thereof.[34] A possible world, on this view, may be a set of propositions or a set of sentences. Since we are likely to believe that there are, at least in some sense, propositions or sets of sentences, the doubts about the existence of possible worlds would be alleviated and we might get an explanation of necessity and possibility out of (nec) and (pos) after all.

Since possible worlds on these accounts are supposed to be some kind of representation, the relevant notion of truth will not be regular truth at a world, but rather something like truth according to the world-story. To simplify, we may take worlds to be sets of propositions and truth at a world to be membership in that world-representing set. To make this difference in what it is to be true at a world explicit, (nec) and (pos) can be modified in the following way:

(nec*) necessarily p, iff p is a member of every world-making set of propositions

(pos*) possibly p, iff p is a member of a world-making set of propositions[35]

The biconditional should, as in the case of genuine modal realism, give us a recipe for giving an explanation of what it takes to be necessarily true for any necessarily true proposition p. The explanation would accordingly run as follows: p is necessary, because it is a member of every world-making set of propositions.

---

[34] Most famously Adams (1974) and Plantinga (1978) . Arguably, these writers are not committed to a non-primitivist version of ersatzism and so may not be subject to the worries outlined here.

[35] I borrow the phrase 'world-making' from Lewis (1986) p. 142 ff.

Once again care must be taken not to trivialize the explanation, but the way in which the explanation should be understood has been outlined above.

At first glance, this may seem like a fairly good explanation. Surely there are propositions (or if you do not like propositions, there are sentences or other representations) and so there are quite a few sets of them. So there is no worry, as in the case of genuine modal realism, that the relevant sets may not turn out to exist. However, there is another problem, which, if it is to be solved, threatens to make the explanation circular: There are sets of propositions which clearly do not represent possibilities. Given that both the proposition that it is raining here right now and the proposition that it is not raining here right now exist, a set containing both propositions exist. But it cannot be that both form a world-making set or are members of a world-making set, for it is not possible that it is both raining and not raining. So while it may be uncontroversial that there are enough sets of propositions, there are too many and we need a way of identifying the ones that are world-making, that is, the ones that represent genuine possibilities. The trouble is that there seems to be no general way to do that which does not itself invoke what is possible.

We may take the world-making sets to be those that consist only of propositions whose conjunction is possible. But in that case, we will get a circular explanation for the propositions which are such conjunctions: A proposition is possible, because it is a member of a world-making set of propositions, and a proposition is a member of a set of world making propositions, because the propositions jointly represent what is possible. Here no explanation of what it takes to be possible is achieved, the explanation is circular.

There are ways of cutting down on the sets of propositions that do not obviously lead to circularity. One may, for example, take the world-making sets to be maximal sets of propositions, that is, sets of propositions that only have constituents that either contain a proposition or its negation, thus avoiding overt contradictions. But this will not suffice, as not only logical contradictions are impossible. It is, for example, also impossible for something to be both a bachelor and married, or to be a closed three sided figure in euclidian space and being a square.[36]

One may still insist that there are ways to identify the world-making sets of propositions without overtly relying on necessity and possibility. This may work. One may, for example, specify the world-making sets as those that are consistent, where consistency is not itself explained in terms of necessity and possibility. But this will leave us with the primitive notion of consistency that seems just as much in need of explanation as necessity and possibility, and so it does not fare well with respect to the intelligibility-requirement on explanation. This option will be further discussed in the section on primitivism, so let me put it aside for the moment.[37]

It is instructive so see why genuine modal realism does not have a problem of circularity while reductive ersatzism does. The modal realist can help himself to possible worlds, because what exists is not itself a modal matter. The space of possible worlds may turn

---

[36] See Lewis (1986) for an extensive discussion of the problems of ersatzism and especially the circularity worry. The version of ersatzism I outlined here comes closest to the linguistic ersatzism Lewis critiques.

[37] Keefe (2002) explores ways in which circularity may not be harmful in philosophical explanations. She explicitly takes a certain ersatzist-view to be an example of a non-harmfully circular philosophical analysis. However, she takes the notion to incompatibility as primitive in her example. Thus subscribing to a version of primitivism, whose merits I discuss in the following section.

out to be very different from what we take to be possible. There may be 17 worlds, or there may be just one. That there is just one is of course what we would ordinarily think, and thus one is inclined to reject the biconditionals (nec) and (pos). Lewis disagrees and thinks that a pluriverse of possible worlds really exists. Importantly, however, both hypotheses may turn out to be the case independently of what we take to be possible.

The situation for the ersatzist is different, as there are no worries about there not being enough sets of propositions. Rather, the worry is that there is no non-circular way to pick out a privileged portion of them, the world-making sets, that have the nice feature of representing only what is possible. And so the problem is not that the analysis in terms of (nec*) and (pos*) is incorrect given our opinion of the truth of the biconditionals, but rather that there is no way of spelling out what a world-making set is that doesn't make the explanation circular and thus unacceptable as an explanation of necessity.

To sum up, while ersatzism does not require us to start believing in many concrete possible worlds, it has difficulties gaining explanatory currency from the biconditionals (nec*) and (pos*), as the explanation is likely to be circular. This is the most important reason why most ersatzers will likely opt for some version of primitivism.

3. Primitivism

The primitivist about modality is dissatisfied both with Lewis' modal realism, and with the reductive ersatzist way of finding substitutes for possible worlds. So while she may still accept the possible-world biconditionals, she does not think that they are to

be used in an explanation of modality.[38] Rather, it is to be accepted that some (broadly) modal notions just cannot be further explained. The task is then to find one such notion that is as intelligible as possible and to show that all other modal notions can be explained in terms of it.

Primitivism, I want to point out, is something of a misnomer, as every philosophical explanation takes some notions as primitive, that is, as understood in the context of the attempted explanation. The label 'primitivism' as opposed to 'reductionism' derives from an unnecessarily restrictive view of philosophical methodology, where conceptual reduction is seen as the only method. In this framework it makes sense to claim that no reformulation of necessity claims is possible in terms of concepts that are not part of some suitably delimited class of modal notions.[39] That no such reformulation is possible, however, is not a sign that the explanation is defective, or somehow worse off than a reductive explanation. For it may just be a contingent fact about our conceptual framework, or about our vocabulary that no such reformulation is possible. Once we take a broader view of philosophical explanation as constitutive explanation, whether a conceptual reduction is possible or whether only an explanation with the help of other modal notion is possible, is not relevant to an assessment of the quality of the explanation. What is relevant, however, is whether the explanans itself is well understood,

---

[38] Lewis (1986) classifies Primitivism as a kind of ersatzism: Magical ersatzism. I prefer to keep genuine ersatzism distinct from primitivism, for the latter explicitly rejects the need for an explanation of at least some modal notions, and is broader than a primitivism based merely on possible worlds.

[39] What counts as a "modal notion" is somewhat unclear. The present debate, however, does not depend on having a clear idea of what counts as a modal notion.

regardless of its status as 'modal notion'. This is what the primitivist has to watch out for, and it is one of the key weaknesses of primitivist theories of necessity. To get a flavor of some such proposals, let me present two.

Robert Stalnaker suggests to think of possible worlds as ways the world might be, where "ways the world might be" are properties of the actual world and are taken as primitive in an explanation of necessity. He writes: "The moderate realist [Stalnaker himself] believes that the only possible worlds there are—ways things might have been—are (like everything that exists at all) elements of our actual world."[40] Ways the world might be are properties of the actual world and so what it takes for something to be possible is that it is a way the world might be. Other writers accept different primitives. Jennifer Wang for example takes the notion of incompatibility between properties to be her primitive modal notion and shows how we can use it to elucidate what is possible and what is necessary.[41] A proposition, according to her, is possibly true, because the properties it predicates are not incompatible properties of something.

Different primitivist proposals may seem like a fairly good philosophical move vis-à-vis genuine modal realism or the circular attempts at ersatzist reduction, but its explanatory merits depend on whether we have a good idea of what it takes for the primitive of the theory to obtain. So, as far as possible, there should be no further interesting question to be asked about what it takes for the primitive to obtain. It should, for example, be clear without further explanation what it takes for two properties to be incompatible. Or it should be clear without further explanation

[40] Stalnaker (2003) p. 32
[41] Wang (2013a), (2013b)

what it takes for something to be a way the world might be. I have my doubts that the particular notions suggested here really move us closer to a constitutive account of what it takes to be necessarily true, but there may be other explanations, broadly within the primitivist camp, which may do a better job.[42]

Since in the following a theory of necessity will be suggested that does not have to take modal notions as primitive, I will contend myself with these remarks for now. What I will have to show, is that the theory I develop scores higher in terms of intelligibility. That is, it should be clearer what it takes for the primitives of the theory to obtain.

4. Essentialism

Using essences as a way of explaining necessary truth has recently gained some popularity. This is especially due to Kit Fine's counterexamples to the standard modal analysis of essence. The standard analysis says that the essential properties of an object just are the ones it has in every possible world in which it exists. However, it seems that not every property an object necessarily has, should count as part of the essence of an object. Consider Socrates and the singleton set containing only Socrates. It is both necessary that if the singleton of Socrates exists, then it has Socrates as a member, and that if Socrates exists, then he is the member of the singleton containing him. But intuitively it is not essential to Socrates to be a member of the singleton containing him, just as it isn't part of Socrates' essence to be such that triangles in euclidian space have an angle-sum of 180 degrees, even though both are properties Socrates has necessarily.

[42] A recently popular line of argument is to ground modality in dispositions or potentialities. See for example Borghini and Williams (2008), and Vetter (2015).

Consequently the necessary properties of an object cannot be identified with its essential properties.[43]

Even if these counterexamples may be countered by restricting the kinds of properties allowed in a specification of essence, they are suggestive of a different direction of explanation running not from modality to essence, but from essence to modality.[44] The explanation suggested by Kit Fine runs as follows: Necessary truths are necessary by being true in virtue of the essence of the objects they are about.[45] It is, for example, necessary that Socrates is human (if he exists), because it is part of Socrates' essence to be human. It is necessarily true that p or not p, because it is part of the essence of the combination of logical connectives that they yield the truth value 'true' if the input is the same proposition. Similar explanations can be given for other necessary truths.

The essentialist theory gives a noncircular and general explanation of what it takes for a truth to be necessary. The non-circularity condition is satisfied, if essences themselves are not spelled out in terms of necessity. The explanation runs, schematically, as follows: p is necessarily true, because its truth flows from the essence of objects (x,y…). There is no obvious reference to modality in the explanation, and so to guarantee non-circularity, the essentialist is committed to not giving an explanation of what it is to flow from the essence of something in modal terms. Since that was the advertised theory in the first place, it is fairly certain that the non-circularity requirement is satisfied.

---

[43] Fine (1994) p. 4ff.

[44] For a suggestion along these lines see Wildman (2013), for another defense of the modalist position see Paul (2006)

[45] Fine (1994) p. 9. Other philosophers defending an essentialist theory of modality are Lowe (2008b) and Hale (2013)

The generality-desideratum is also likely to be satisfied, provided that essences are rich enough to cover all necessary truths in need of explanation. So in principle at least, the explanation of necessity in terms of essence should turn out to be general.[46]

The trouble with the explanation of necessity in terms of essence lies, as so often, with the last desideratum: intelligibility. The trouble is twofold. First, it seems that essences themselves are in need of explanation. Second, some clarification of what it takes for a truth to flow from the essence of something or what it is to be true in virtue of essence would be desirable. These two points will be elaborated on in chapter 8 of part two, so presently I will be content with just giving an outline of the two potential problems.

The first problem comes out best when considering why philosophers have taken essence to be explicable in terms of necessity. The reason was philosophical dissatisfaction with the concept of essence itself and the reasonable hope that an explanation of modality may be easier to come by than an explanation of essence. While these hopes may not be justified, given the theories of necessity surveyed here and the apparent counterexamples to such analyses, it still remains to be seen whether a good explanation of what it takes to be (part of) the essence of something can be given. Some attempts have been made to think of essence as given by a real definition, but, as before, it is at least unclear that this helps to clarify the notion.[47]

---

[46] Fine (2002) himself has some doubts about it and seems to believe that some other kinds of modality, for example, natural necessity, should be accounted for differently. Some of these intricacies are discussed in Part II, Chapter 8 of the present thesis.

[47] see Fine (1994), and Lowe (2001) for attempts to account for essence in terms of real definition.

The second problem amounts to the question of how the fact that something is (part of) the essence of something gives rise to the necessary truth of a proposition involving the object which has this essence. Why, for example, should the fact that Socrates is essentially human, make the proposition that Socrates is human necessarily true? The answer an essentialist may be prone to give is the following: being human is part of what it is to be Socrates, and so Socrates cannot exist without being human. So it is necessarily the case that he is human, if he exists. The trouble with this answer is that being part of the essence of something is identified with being a necessary condition for an object's existence. But if essence just is a necessary condition for the existence of Socrates, then essence is again explained in modal terms, which the essentialist cannot do on pain of circularity. So, just as in the case of primitivism, it is unclear what it takes for the explanans, being true in virtue of essence, to be the case. The challenge for the essentialist is therefore to give some answer to what it takes for a truth to be "true in virtue of essence" and how this gives rise to necessity.

All this is not to say that there is no way to make this theory work, and even a certain primitivism about truth in virtue of essence may be a viable fallback position, but it highlights the downside of a theory of necessity in terms of essence.

Summing up

I outlined four prominent ways of giving a theory of modality and discussed some of their drawbacks as well as strengths. I did not give a knockdown-arguments against any of the positions, but some general understanding of the difficulties of giving a theory of modality could, I hope, be gained: Reductive possible world theories face the problem that they need to give a non-circular

account of what possible worlds are. The only account which succeeds in this endeavor, genuine modal realism, either requires us to believe in a pluriverse of other concrete worlds, or yields that nothing non-actual is possible. Both options seem a high price to pay. The way out, taking uncontroversial ersatz-entites to play the role of possible worlds, fails, because the only way to delimit the ersatz-entities is by recourse to possibility and is thus circular. These findings may push one towards accepting some primitivism about modality and thereby incur a certain risk that the primitives of the theory are themselves in need of explanation. Proponents of essentialism incur a similar risk. For its basic explanatory notion, truth in virtue of essence, may itself be in need of further explanation.

So it is safe to say that the biggest challenge for a viable theory of modality is that the explanans itself must be intelligible, such that there is no interesting further constitutive question about what it takes for it to obtain.[48]

[48] I do not want to exclude the possibility that there is always some such further question to be asked. Indeed it is quite likely that there is always some such further question to be asked. However, these questions should be philosophically less interesting than the original question.

Chapter 3: Analyticity and Necessity

An explanation of necessity that was omitted in the above list is the once popular, now scarcely defended view that the necessary truths just are the analytic truths. In the following this theory and some of its problems will be discussed. The difficulties of such a theory, I argue, justify today's almost universal contempt for it. However, something insightful can nonetheless be gained from some of the roots of these theories. In particular, the conception of analyticity advanced by Kant will be shown to provide us with a partial blueprint of how a successful explanation of necessity may be effected.

1. Necessity as Analyticity

Not so long ago analyticity, necessity, and *a priority* were used almost interchangeably: Whatever is a necessary truth is also analytic and can be known *a priori*.[49] Or so the consensus went. The reasons for adopting this position were typically a strong commitment to the empiricist principles of logical positivism, but can also be motivated without the theoretical baggage: Many standard examples of necessary truths are either tautologies (consider: p then p, p or not p), or follow from definitions, (consider: bachelors are unmarried men, knowledge is justified true belief, etc). Since analytic truths are, on the view abstracted from Frege, transformable into logical truths with the help of

[49] See for example Ayer (1936a) and (1936b), but also Carnap (1947) and (1952).

definitions, these examples should all count as analytic.[50] And since logical truths are uncontroversially knowable *a priori*, and the definitions are just an analysis of meaning, which can be carried out in an *a priori* manner as well, the necessary and analytic truths are all knowable *a priori*. The converse direction is also supposed to hold. The reasoning for this is not as straightforward, but assuming that *a priori* conceptual analysis is the only philosophical method which provides a way of arriving at necessary truths, it is natural to suppose that all three concepts are coextensive. The apparent fact that all analytic truths are necessary as well as knowable *a priori* still leaves open the explanatory relations between these properties of sentences or propositions. But a certain picture suggests itself: both necessity and *a priority* may be explained by analyticity. The reasoning for this proceeds from the dual observation that analyticity is truth in virtue of meaning, and that meaning is a matter of convention. With analyticity as truth by convention we get an explanation of the necessity as well as *a priority* of the truths that are true by convention: They are *a priori*, because conventions are something that everyone who is party to the group in which the convention holds knows and so does not have to figure out from observation. And they are necessary, because their truth is required by the

[50] Frege, (1884) GLA §3 p. 14-15 writes something slightly different: "Es kommt nun darauf an, den Beweis zu finden und ihn bis auf die Urwahrheit zurückzuverfolgen. Stößt man auf diesem Wege nur auf die allgemeinen logischen Gesetze und Definitionen, so hat man eine analytische Wahrheit, wobei vorausgesetzt wird, dass auch die Sätze mit in Betracht gezogen werden, auf denen etwa die Zulässigkeit einer Definition beruht." His concern is thus not so much with transformability into logical truths, but with the ingredients of a proof of the analytic truth.

conventions, if the convention is constitutive of the meaning of the true sentence.[51]

At a time, this view may have seemed so obvious, that it hardly needed spelling out, but from today's perspective it seems, to say the least, confused. This is for two main reasons. First, it seems that there are clear counterexamples to the claim that analyticity, necessity, and *a priority* are coextensive. Second, it is unclear how some truth-bearer can be true by convention, even if it is admitted that meaning was purely a matter of convention. The counterexamples to necessary truths being *a priori* show that there is a need to at least supplement the account to be able to accommodate those necessary truths which cannot be known *a priori*. The second criticism is stronger, because it shows that the underlying concept of truth by convention is in need of serious reevaluation. I will go over both objections in turn.[52]

The stock examples for necessary *a posteriori* truths come from Kripke and Putnam. Kripke asks us to consider the identity statement "Hesperus is Phosphorus". Since identities are necessarily true, if the names are rigid designators, it is necessarily true that Hesperus is Phosphorus.[53] But, as is well known in philosophical circles since Frege, it is not *a priori*

---

[51] Here is what Ayer (1936a) p. 20 has to say about it: "They [the analytic and *a priori* propositions] make no statement whose truth can be accepted or denied. They merely lay down a rule which can be followed or disobeyed."

[52] Historically, Quine's criticism of the very distinction between analytic and synthetic sentences was very influential. I will, however, not discuss Quine's points explicitly here, as the points I do discuss partly incorporate them. Where they don't, Quine's criticism focusses largely on the idiosyncrasies of his specific opponent, and are thus only of marginal interest to the general project of assessing the merits of a theory of necessity in terms of analyticity. The interested reader may consult the following works by Quine for his attacks on the distinction: Quine (1951a), Quine (1960), Quine (1976).

[53] Kripke (1980)

knowable that Hesperus is Phosphorus, as it was a quite substantial empirical discovery that both names denote the same planet, that is, Venus.[54] Another example comes from Putnam, who shows with the help of his twin-earth thought experiment that water necessarily has its molecular structure (H_2O).[55] Obviously, the molecular structure of water was discovered and can only be discovered *a posteriori*. Both examples show that necessity and *a priority* can come apart, and if analytic truths are considered to be *a priori* on the grounds that they are true by convention, then the examples show that not all necessary truths are true by convention. This constitutes a significant blow to the view that necessity can be explained by analyticity. At least it shows that the explanation cannot work in the simple way envisaged above.

There have been attempts at modifying the original theory to accommodate the counterexamples. Alan Sidelle, for example, argues that even though these necessary *a posteriori* truths are *a posteriori*, their necessity still derives from an analytical core plus some non-modal *a posteriori* premises.[56]

The idea utilizes some of the insights that have been gained in the study of meaning. Consider indexical expressions like ‘I’, ‘here’, and ‘now’. ‘I’ when uttered by me, denotes me, when uttered by you, it denotes you. So on different occasions of use, in a different context, a different referent is determined. Still, in some sense, ‘I’ when uttered by me, has the same meaning as ‘I’, when uttered by you. This core meaning, or character, as it is commonly called following Kaplan, stays the same. Formally, character can then be

---

[54] Frege (1948)
[55] Putnam (1973) and (1975)
[56] Sidelle (1989)

viewed as a function or rule from a context to a referent, and may plausibly be taken to be a matter of convention.[57] According to the causal theory of reference advanced by Putnam and Kripke, which was designed to account for the reference of names or natural kind terms, names and natural kind terms work similarly: While there is no change of reference with changes in context, we can nonetheless split up what determines the reference into a conventional part and an external part. The conventional part says, for example that the name is to refer to the person to which it has been given, while it is fixed externally who it was given to. Similarly in the case of water: the conventional part of what determines reference is that the natural kind term water is to refer to that substance which was named water. Which substance was named water is again fixed externally, it just so happens that it was actually H_2O. We can think of the core meaning, or character, in these cases as a function from a context of introduction to a referent, and also view it as a matter of convention.[58]

Sidelle's idea is that we may admit that it is not *a priori* knowable that water is H_2O, because we do not know what the substance that was named 'water' is. Similarly in the case of "Hesperus is Phosphorus". But we do know the core-meaning of the terms *a priori*, since it is a matter of convention, and so at least the following conditional is knowable *a priori*: If the substance at the name-giving ceremony was H_2O, then necessarily, water is H_2O. Thus we come to know, by finding out that the substance is indeed H_2O, that water is necessarily H_2O. But the necessity, Sidelle suggests, is explained by the conventions, because the

---

[57] For the still standard treatment of indexicals see Kaplan (1977).

[58] Gillian Russell (2008) introduces this terminology as well as a more detailed formal treatment.

only *a posteriori* element in arriving at the necessary truth that water is H_2O, is the non-modal premise that H_2O is the substance indicated at the name-giving ceremony. The premise involving the necessity, however, is knowable *a priori*, and thus we should conclude that necessity is explained by convention.

The suggestion of splitting up the external factors and the internal factors of reference determination was taken up by philosophers of language as well as defenders of conceptual analysis, and now often runs under the broad label of two dimensional semantics.[59] While the insight that reference determination of words works via conventional and non-conventional elements is unobjectionable, even if contestable in the individual case, Sidelle's suggestion does not quite show that we should conceive of necessity as being a matter of conventional truth, for it doesn't follow from the fact that we can know the conditional premise *a priori*, that its truth is merely a matter of convention.[60]

The last point is an instance of the second problem of conventionalist theories of necessity: How can conventions possibly make a sentence true? Surely it is at least partly a conventional matter what the content of some sentence is. But that seems to be all linguistic conventions can do: they tell us what proposition is expressed by a sentence. They cannot also do the job of making what is expressed the case. Here is how Paul Boghossian puts the point succinctly: "Isn't it generally true - indeed isn't it in general a truism - that for any statement S, S is true iff for some p, S means that p and p. How could the mere fact

---

[59] see for example: Jackson (2000), Chalmers (2006), Russell (2008). Chalmers and Jackson use the two-dimensional framework explicitly to account for the epistemology of modality.

[60] This point is forcefully made by Yablo (1992) in his review of Sidelle's book. He attributes the basic point to Lewy (1976)

that S means that p make it the case that S is true? Doesn't it also have to be the case that p?"[61] Just as it is partly a matter of convention and partly a matter of what is the case that the sentence "the sun is shining" is true, it is partly a matter of convention and partly a matter of what is the case that the sentence "all bachelors are unmarried" is true. This rather obvious observation is a strong argument against the conventionalist thesis. If it has to be the case for the sentence "all bachelors are unmarried" to be true that really all bachelors are unmarried, then it is false that its truth is due to its meaning, for it is due to the world being such as the sentence says it is. This observation shows that no sentence, not even the ones we would want to count as analytic, are true by convention, or as it is sometimes put, true in virtue of meaning.

It also reveals a confusion by the conventionalist about what the objects that are said to be necessarily true are. For what is necessarily true is not a sentence, but what is expressed by the sentence. At the very least it is an interpreted sentence, a sentence with its full meaning. Assume, for reductio, that it is sentences that are necessarily true. Assume the sentence "All bachelors are men" was written on the wall of Humboldt University. As of now,

---

[61] Boghossian (1996), p. 364. For similar points see also: Yablo (1992), Lewy (1976), Harman (1996), Salmon (1993), Hale (2002) p.305/p.306. Williamson (2006), p.8 footnote 8 puts it this way: "It is widely acknowledged that 'Vixens are female foxes' is true not simply because it means that vixens are female foxes but because it means that vixens are female foxes and vixens are female foxes, just as 'Vixens are hunted' is true not simply because it means that vixens are hunted but because it means that vixens are hunted and vixens are hunted. Of course, it is necessary that every sentence which means that vixens are female foxes is true, and merely contingent whether every sentence which means that vixens are hunted is true, but that does not show that 'Vixens are female foxes' is true in virtue of meaning in any interesting sense, [...], in the absence of an independently established connection between meaning and necessity."

this expresses a necessary truth. But imagine that the use of the word 'bachelor' changes to not only include men, but more people on the gender-spectrum. Then the sentence as written on the wall will turn out to be false at some point. Consequently, it cannot be the uninterpreted sentence that is necessarily true, but must be the sentence *cum* meaning. This admittedly rather trivial point goes to show that the object which is in the business of being true or false, and so necessarily true or false, is not the sentence independently of its meaning, but only the fully interpreted sentence.

The sense in which a sentence is (at least partly) subject to conventions, in contrast, is in the uninterpreted sense. It may, for example, be that some sentence (in the uninterpreted sense) is subject to a convention that it is always to be interpreted such as to turn out true. While such a convention may render the sentence (at least as long as the convention applies) true, it would do so only by appropriately adapting the interpretation such that it always says something that is true. Real examples of such cases, with less explicit conventions, exist for sentences with only partial conventional meaning: Consider the sentence "I am here now". This sentence, whenever uttered by someone, turns out true, because the contextual features for completing the interpretation of the sentence are to be selected such that 'I' denotes the speaker of the context, 'here' denotes the place of the (speaker of) the context, and 'now' denotes the time of the utterance.[62] Consequently, whenever uttered, the sentence will

[62] This is a simplification, for there are cases in which this does not hold. Consider, for example, answering machines, where the line "I am not here right now" is both quite common and usually true. Such cases are discussed under the label of "the answering-machine paradox". See for example Sidelle (1991). The simplification will suffice for the moment, however.

express a truth. However, just because this sentence is such that the conventions render it true whenever uttered, what it expresses is not necessarily true. To the contrary, what it expresses at each utterance is a contingent truth: I could have been at another place than I am now.[63]

This shows that the conventionalists have confused the objects that can properly be said to be necessarily true and the objects that are subject to conventions. This confusion in turn leads to a confusion between two different properties, which apply to two different kinds of objects: truth by convention, which applies to sentences, and necessary truth, which applies to representations with a full meaning or to propositions.[64] Since some sentences that are true by convention can be representations that are not necessarily true, a conventionalist theory of necessary truth is bound to fail.

This is all rather well known territory in the current philosophical discussion, but there have recently been some suggestions as to how the concept of analyticity, conceived of as a property of sentences, can nonetheless play some role as respectable philosophical concept, and how necessity may nonetheless be explained as arising from conventions.

---

[63] Gillian Russell (2010) recently made a similar point in a paper, arguing that conventionalism about necessity would have to wrongly classify sentences like "I am here" as necessarily true. She does not offer the same diagnosis though.

[64] García-Carpintero and Pérez Otero (2009) offer a similar diagnosis in their defense of analyticity against Boghossian's argument. They write (p. 240-241): "The probem with it [arguments like Boghossian's] as a refutation of that view [that there are analytic truths] lies in that it assumes a dichotomy of candidates for the role of truth-bearer that is not exhaustive. It assumes two kinds of bearers of truth and modally qualified truth: on the one hand, linguistic items individuated merely by their "formal" (phonological or graphic, and syntactical) properties; on the other, Platonistic propositions - propositions whose nature and properties do not depend on facts about thinkers, in particular facts about the intentions and conventions guiding them."

The first line of reasoning is adopted by Gillian Russell, who updates the concept of analyticity and argues that it should play a role in epistemology distinguished from *a priori* knowledge.[65] However, she explicitly rejects that this conception has any role to play in an explanation of necessary truth. Russell's basic idea derives from the suggestion that we can sometimes split up the factors contributing to the content of sentences into different factors, for example into character plus context, which only together determine a reference. Russell's suggestion is that this picture of language is still too simple. Rather we should not only allow for traditional context-sensitivity and the correlative notion of character, but also acknowledge that reference determination may be sensitive to what she calls a context of introduction of a word. This idea derives from the causal theories of reference advocated by Kripke and Putnam, where the situation at the context of introduction of a word plays an important role in determining its reference. Analytic truths, Russell argues, are then to be understood as those which cannot fail to express some truth, no matter how the context of introduction and the context of utterance are varied. Here is how she puts it: "A sentence S is true in virtue of meaning just in case for all pairs of context of introduction and context of utterance, the proposition expressed by S with respect to those contexts is true in the context of evaluation."[66]

---

[65] Russell (2008)

[66] Russell (2008) p. 56. Russell also gives a further definition not depending on the modal framework, in terms of containment of what she calls reference determiners (2008, p. 100). Reference determiners should be thought of as the core conditions that have to be satisfied for reference of a word to be determined, and may need to be supplemented by the context of utterance or context of introduction to determine a reference.

This definition deliberately makes analyticity, or truth in virtue of meaning, a property of only partially interpreted sentences. What is expressed by these sentences need not be necessarily true, such that contingent truth like "I am here" (at least under simplified assumptions) turn out to be analytic. Russell claims that even though this notion of analyticity explains neither necessary truth, nor the *a priori*, it has interesting epistemic consequences. However strong the last claim may turn out to be, by explicitly making analyticity a property of sentences, and denying that analyticity explains necessary truth, the account escapes the criticism leveled against more traditional accounts of truth in virtue of meaning. I will put the account aside for the present discussion, however, as the current project is the explanation of necessary truth, for which a different conception of analyticity is needed.

A different answer to the challenge that conventions are not what makes propositions true consists in taking a more globally anti-realist approach to necessity. Answers along these lines are advocated by Cameron, Sider, and in more detail by Thomasson.[67] Cameron argues that instead of taking truth to be a matter of convention, we may take necessity itself to be a matter of convention: The way we divide up all truths into the necessary and contingent is itself a matter of convention. Cameron invokes an analogy to elucidate the view: Just like whether a city is on one

[67] Cameron (2010b), Thomasson (2013) and a forthcoming book. Ted Sider (2003) usefully distinguishes between a "governance" and a "classification" conception of necessity. On the governance conception the source of the truth itself lies in the (source of the) necessity of the truth. Truth in virtue of meaning would be an example of this kind of conception. On the classification conception, however, truth comes first and its necessity is explained by something else. Sider suggests that while the governance conception is untenable, a classification conception can possibly be worked out.

side of a border or on the other is, in a sense, just a matter of convention, of where we choose to draw the border, whether a truth falls on the necessity-side or the contingency-side of the necessary-contingent divide, is a matter of where we draw the boundary.[68] This does not mean that it is arbitrary or that there is no fact of the matter on which side of the border the city is, or whether a truth is necessary or contingent, but it is in virtue of complex social facts that some truths *count as* necessary or contingent respectively. What these facts are, Cameron suggests, can be found out by investigating why we use modal idiom.

Thomasson similarly argues, that we should take our attributions of necessity to not have their ground in any special way of being true, but rather in our using necessary truths for normative purposes. More specifically, we use modal idiom to express rules of language. To support her thesis, Thomasson shows in great detail how we can move from expressions of semantic rules to modal-attributions and back and cites a range of evidence of where modal idiom is used in just this way.

Some, but possibly solvable, problems for such accounts remain, for example, how to account for *de re* modality and other kinds of necessity such as nomic necessity. But these difficulties shall not concern us at present. I merely note that even if it can be convincingly shown that modal talk is sometimes, or even often, used in normative contexts, for example to point to rules of language, it does not quite establish the truth of the neo-conventionalist claims. For the claim cannot only be that they are thus used, but that this use is constitutive of what 'necessarily' means. This, however, should be taken to be true only if there is

[68] Cameron (2010b) p. 355

no strong reason to believe that there is an explanation of necessity not dependent on us using some (true) sentences in the suggested normative way.

Since I will argue in the following that there is very good reason to believe that some truths are substantially necessarily true without having to be used in a normative way, I take it that these theories, while telling us something interesting about how we can make use of necessary truths in certain normative contexts, may fail in their more global ambitions to constitutively explain what necessary truth is.

2. Analyticity without Convention

Leaving aside the normativist accounts outlined above, explanations of necessity in terms of analyticity are problematic, because they appeal to truth by convention and conventions, while plausibly involved in giving meaning to words and sentences, cannot make true what a sentence says. However, thinking of the property of analyticity as applying to sentences, and to take these analytic sentences to be true by convention, is not the only way of conceiving of analyticity. It certainly is not the way in which Kant, probably the source of much talk of analyticity thereafter, thought of it. In what follows, I will show how we may think of analyticity without conventions, by moving from analyticity as a property of sentences, to analyticity as a property of representations. Analyticity, in this broadly Kantian spirit, should thus not be taken to be truth in virtue of meaning. Nonetheless, the analytic truths in the Kantian sense can plausibly be taken to be necessarily true, as well as necessary in virtue of meaning (in a sense to be qualified).

The Kantian account of analyticity does not start from words and sentences, but from concepts and judgments. This makes the account non-conventionalist from the start. For while it may seem plausible that the meaning of words and sentences is contingently associated with them by convention, the content of concepts and judgments is essential to the concept and judgment itself.

Kant provides the following definition of analyticity as a property of judgments, a definition in terms of concept containment:

"In all judgments in which the relation of a subject to the predicate is thought […] this relation is possible in two different ways. Either the predicate *B* belongs to the subject *A* as something that is (covertly) contained in this concept *A*; or *B* lies entirely outside the concept *A*, though to be sure it stands in connection with it. In the first case, I call the judgment analytic, in the second synthetic."[69]

Note first that there is no mention, neither explicitly nor implicitly, of truth or even meaning at all. The objects that are said to be analytic are judgments that relate a predicate to a subject. Judgments of this kind, it seems to be assumed, are composed of concepts and concepts themselves may be composed of other concepts.

What exactly judgments as well as concepts are, according to Kant, need not be discussed in detail.[70] For the account to work,

---

[69] Kant KrV A 6-7: Translation from Kant (1998).

[70] The present remarks are not intended to give an exegetically grounded analysis of Kant's conception of judgments. I merely note some features of judgments and their constituents presupposed by the definition of analyticity. Also I do not want to claim that the account of representations developed later is Kantian in any sensible way. Rather, the conception of judgments and concepts presupposed by Kant's definition of analyticity has some interesting features that differ markedly from newer accounts of analyticity, which take it to be a property of sentences.

however, some general features of judgments and concepts need to be presupposed. First, judgments and concepts are not to be conceived of as just sentences and words. We should rather think of them as types of representations, which have their representational features essentially. This allows for them to be types of concrete thoughts or statements that are individuated with respect to what they represent. How finely we should individuate judgments and concepts is an intricate matter and there is no need to assume any specific view at the moment. Concepts can be thought of as building blocks of both judgments as well as the building blocks of more complex concepts, but the mode in which they contribute to the judgment and the complex concept respectively is different. In judgments, the way in which the concepts come together determines what it takes for that judgment to be true: In the case of subject-predicate constructions the referent(s) of the subject concept must be a subset of the referents of the predicate concept. So by constituting the judgment, the referents of the concepts are said to stand in a certain relation to each other. In complex concepts, the constituent concepts have a slightly different (if related) function. Here the reference of the complex concept depends on the reference of the constituent concepts. In the standard case, the reference of the complex concept will be determined to be the intersection of its constituent concepts. To use Kant's own example, consider the concept 'bachelor', which is said to consist of the concepts 'unmarried'

and ‘male’ and thereby has its reference determined by the concepts ‘unmarried’ and ‘male’ to be their intersection.[71]

With this background, the broadly Kantian definition of analyticity comes quite naturally: Some judgments may say of things which are A that they are B, where B is a constituent of the concept of A. Analytic judgments are thus very literally analytic. One can also give a recipe for constructing them: To arrive at an analytic truth decompose a concept (for example A into B and C), and form a judgment from the complex concept and (one, or both of) its constituents (A is B, A is C, A is B and C).

How does all this relate to necessity? The connection is fairly straightforward, and has to do with the way the simple concepts play their role both in the constitution of the judgment, as well as in the constitution of the complex concept. The judgement’s truth requires a certain relationship between the referents of its constituent concepts. In the case of the simple subject-predicate judgment, the requirement for its truth is that the referent(s) of the subject concept are a subset of the referents of the predicate concept.[72] In the case of analytic judgments the predicate concept is contained in the subject concept. So the reference of the subject concept is determined partly by the predicate concept. Specifically, the reference of the subject concept is determined

---

[71] Concept constitution is problematic, for it is not clear that constitution always gives rise to the referential dependencies outlined here. Presently we may circumvent this problems, by taking concept constitution proper to be that kind of constitution which gives rise to the relevant referential dependencies.

[72] For the moment the case of non-referring subject concepts is neglected. We may generally take such judgments to be true, as seems reasonable in the case of bachelor. This is problematic in the case of singular subject concepts. In modern discussions, a more complex logical form of the judgment solves the problem. But presently it is also not objectionable to have different requirements for truth for different subject predicate constructions, as I do not claim that constitution settles everything about what is required for truth of a judgment.

such that it is a condition on being the referent of the subject concept that it is also a referent of the predicate concept.

By considering the constitution of the concept and the judgment, we gain two facts about the judgment and the concepts constituting it. First, by constitution of the judgment, the subject-predicate judgment is true, just in case the referents of the subject concept are referents of the predicate concept. Second, by concept constitution, the referents of the subject concept are referents of the predicate concept.

From these two facts it can straightforwardly be concluded that the requirements for truth of the judgments are bound to be satisfied, no matter what the actual referents of the concepts are. So we can infer the judgment's necessary truth from considerations about how reference is determined and from which relation between the referents is required for the truth of the judgment. Necessary truth of a judgment thus arises from an appropriate relationship between what is required for the truth of the judgment, and what determines the reference of the constituent concepts.

This is, it seems, both a straightforward explanation of the necessity of a subject-predicate judgment as well as an account of analyticity that does not fall prey to the main objection raised against accounts of analyticity which take it to be truth in virtue of meaning. I will go over both points in turn.

First, consider the explanation of analyticity of "Bachelors are unmarried". Why is this analytic according to the Kantian explanation? Because the predicate concept is contained in the subject concept. Why is it true? Because the referents of 'bachelor' are a subset of the referents of 'unmarried'. Why are such analytic judgments interesting? Because they are necessarily

true in virtue of how the reference of their constituents is determined in relation to what is required for their truth.

These answers do not talk of truth in virtue of meaning at all, or any other special kind of truth for that matter. Meaning, in so far as it is understood as what determines the reference of concepts, as well as what determines the requirements for truth, does play a role in the explanation of analyticity as well as necessity, but it is not invoked in an illicit manner in the explanation of why the judgment is true, for the role 'meaning' plays is just the same role it plays in any explanation of why an ordinary, non-analytic judgment is true. For consider a true non-analytic subject predicate judgment: Cabs are black. Again the judgment's constitution determines what it takes to be true: "Cabs are black" is true, just in case the referents of the concept 'Cabs' are a subset of the referents of the concept 'black'. In contrast to the analyticity case, we have no obvious concept containment relations between the concepts 'Cabs' and 'black', and so nothing that guarantees that the relation between the referents required for the truth holds. Still the explanation for the truth of the judgment is the same: it is true because the referents of 'Cabs' are a subset of the referents of 'black'. The relevant difference between the two cases is just that in the case of the analytic judgment, the reference determination of the constituent concepts was appropriately related and it was related such as to guarantee that the requirements for truth were satisfied.

This finding also provides us with an explanation of why the particular judgment "bachelors are unmarried" is necessarily true: It is necessarily true, because the reference of its constituent concepts is determined such that its requirements for truth, determined by the constitution of the concepts, are guaranteed to

be satisfied. In the present case this guarantee is effected by the containment relations holding between the concepts plus the subject predicate structure of the judgment.

It seems, therefore, that Kant's definition of analyticity, appropriately understood, gives us a reasonable, if narrow, conception of analyticity, as well as a non-circular and illuminating explanation of the necessary truth of some judgments. Why, one may wonder, was this account abandoned in the subsequent discussion on analyticity?

Part of the reason was surely that the notion of concept-containment was perceived as too unclear or metaphorical, especially in light of the formal developments in logic at the end of the nineteeth and beginning of the twentieth century.[73] There is some reason to believe, however, that Kant was using a technical notion of containment, which did have a clearer meaning than is apparent today.[74] More importantly, there is a way of reading the definition that does not take containment literally, but takes seriously the role containment plays in the original definition: guaranteeing a referential dependency between the representations in subject and predicate position. So a minimal understanding of 'containment' as mere referential dependency, should assuage worries about the metaphorical character of containment.

A further factor that may have played a role in rejecting the Kantian definition of analyticity was the shift away from talk of concepts and judgments to words and sentences, a shift that was facilitated by the development of modern logic. But one of the

---

[73] See Frege (1884)

[74] For a recent and detailed, historically rich defense of the concept of containment see Anderson (2015)

main reasons for the demise of the Kantian view of analyticity was probably Kant's own (justified) insistence that analyticity, if conceived of in terms of containment, and concerning only subject-predicate constructions, is not very interesting, because there just are not that many philosophically interesting analytic truths. After giving his definition of analyticity in the Critique of Pure Reason, he goes on to show that a large number of truths which are apparently knowable *a priori*, are not analytic at all. Mathematical truths, for example, do not seem to exhibit any concept-containment, but are nonetheless *a priori* as well as necessarily true. Similar considerations apply in other domains.

So the trouble with the Kantian containment account, despite its merit in explaining some cases of necessary truth, is that it is insufficiently general. Possibly, some simple truths like "all bachelors are unmarried", "vixens are female foxes", truths that contain explicitly composed concepts such as "red balls are red", and truths we can build from explicit definitions do count as analytic in the Kantian sense. But it is quite unlikely that any such account can work for philosophically interesting cases. Thus a general explanation of necessary truth in terms of containment is not to be had.

We need to be careful, however, what this reasoning demonstrates, and what it does not. It shows that concept containment by itself can neither explain *a priori* knowledge nor necessary truth generally. But it does not show that it is not part of a good explanation of necessary truth in the cases where there is concept containment in the relevant sense. Better still, the way in which necessary truth is explained, by an appropriate relationship between what determines reference and what it takes for a representation to be true, may point us to an explanation of

necessity that is much more widely applicable. For the ingredients of the kind of explanation, reference determination of constituent representation together with an appropriate relation to what it takes for a representation to be true, require neither complex concepts, nor subject-predicate judgments, but merely something that determines the reference of constituent representations, and the dependency of truth on the relationship of the referents. This may be found in a much wider range of cases. So while the narrow Kantian containment analyticity may not be widely applicable, a wide conception of analyticity, no more chained to containment and subject-predicate judgments, may be a much more powerful notion.

It is the purpose of the second part to develop the framework in which such explanations become possible. For now, the important lesson is that the definition of analyticity, as given by Kant, does not exhibit the biggest traditional problem of analyticity: the problem that there is no truth purely in virtue of meaning. For analyticity, in the Kantian spirit, is not to be conceived of as truth in virtue of meaning, but as necessity in virtue of reference determination.

Summing up

The best known theories of necessity utilizing analyticity are spelled out in terms of conventional truth. Analytic truths, on this view, are true by convention alone. This special way of being true is also supposed to explain their necessity, as there is no way of taking them to be false which does not require a change in convention. A welcome consequence for the proponents of the account is that the necessary truths may turn out to be knowable *a priori*, if knowledge of conventions can be taken for granted. This

forms a tight connection between analyticity, *a priori* knowledge, and necessity, and thus seems to validate the radical empiricism of some of its proponents.

Conventionalist accounts, however, are unsustainable. Counterexamples to the *a priority* of some necessary truths have shown that one must either give up analyticity's connection to *a priori* knowledge or the connection to necessity. But even worse for the proponents of this theory, it was convincingly argued by a range of philosophers that linguistic convention cannot be what makes a sentence true: no sentence is true in virtue of meaning alone. The most conventions can do is give the sentence a meaning, but for it to be true, what it says needs to be the case.

This seriously undermines the project of explaining necessary truth in term of analyticity. Friends of a conventionalist account of necessity may abandon the concept of analyticity, and seek to make the attribution of necessity a way of expressing conventions. This does not fall victim to the objection, but also makes necessary truth a rather insubstantial matter. Friends of analyticity, on the other hand, may abandon the search for an explanation of necessity and *a priori* knowledge in terms of analyticity, and find some different use for a suitably refined version of analyticity.

Help comes from an unsuspected source: the Kantian containment-conception of analyticity. The Kantian account does not fall victim to the objection that there is no such thing as truth in virtue of meaning, for it does not attempt to account for necessity in terms of truth in virtue of meaning, but rather gives us a way of explaining necessity in virtue of reference-determination: The containment of a concept (A) in another (B) explains that the reference of the complex concept is determined

in such a way as to guarantee that the requirements for truth of the subject-predicate judgment "B is A" are satisfied.

As it stands, this account is much too narrow and, as Kripke's and Putnam's examples have shown, the link between analyticity and *a priori* knowledge needs to be given up, if a general account of necessary truth should be constructed from the basic Kant-inspired idea.[75]

I will devote the second part of this thesis to arguing that necessity generally can be explained in this way, once we move away from an overly narrow conception of what determines reference. The crucial move to make such an explanation possible, however, is inspired by the Kantian account of analyticity. It is to move away from mere sentences as the objects of study, and instead to representations, objects which have their representational properties essentially. Analyticity and necessity should be conceived of as a properties that apply to representations, which hold of truth-evaluable representations in virtue of the appropriate relationship between what determines the referents of the constituents and what is required for the truth of the representation.

If this explanation can be shown to work generally, I take it to provide an illuminating and intelligible explanation of necessary truth. Convincing the reader that this is the case, is the project of the second part of this thesis.

---

[75] I do not want to claim that the resulting account of necessity is Kant's account of necessity. In fact, it probably isn't. What Kant's account of analyticity does for the present theory, is to provide a blueprint of how necessary truth may be explained.

Part II

Chapter 4: Representation, Reference, and Truth

The discussion at the end of the last part suggested that a broadly Kantian conception of analyticity may provide a blueprint for an explanation of necessity. This Kantian conception of analyticity made use of judgments, concepts, and containment relations between them, not of words and sentences and their meanings, as was the case for twentieth-century accounts of analyticity. This difference in the objects to which analyticity and necessary truth attach is crucial to the project of giving an explanation of necessary truth in terms of the relationship between reference determination and what is required for the truth of a representation. To give a fully general account of necessary truth, a framework and new terminology will be developed. The starting point of the account are referential and truth-evaluable representations. Representations, in contrast to words and sentences, have their referential properties essentially.[76] As such they are the proper bearers of necessary truth, or so I will argue.

It is the task of the present chapter to develop a detailed account of representations and their properties relevant to an explanation of necessary truth. To clarify the difference between more orthodox accounts of the semantics of words and sentences, and an account of the relationship between representations and the world, as well as to motivate the move from words and sentences

---

[76] Fine (1994) calls this a "thicker and perhaps more natural conception" of a word, and uses the example of the word bank: On the thin conception this is one word which may either denote the bank for sitting or the bank for money. On the thicker conception the word 'bank' as used for the bank for sitting, is a different word from the word 'bank' as used for the bank for money. García-Carpintero and Pérez Otero (2009) similarly urge us to think of words not just as individuated by their formal properties, but also by their meaning.

to representations, I will first discuss a certain picture of language underlying the classical semantic project: the communication model. This communication model will then be used as a foil to elucidate my own account of representations.

1. The Communication Model

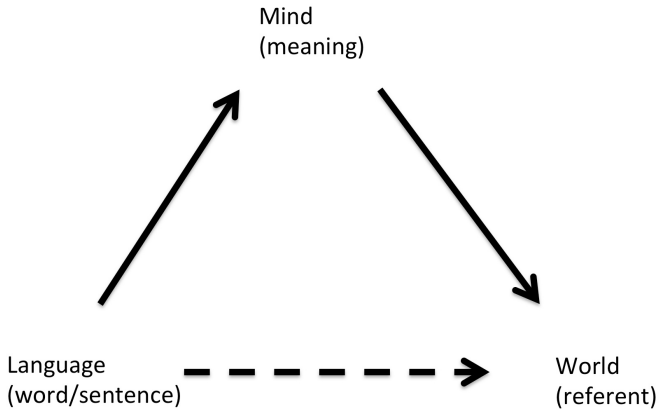
Human beings use language to communicate with each other. Sentences and their constituent words are the vehicles we use to convey information. For words and sentences to play this role in human interaction, they need to have a meaning. This meaning is known by the members of the community speaking that language, and so when a speaker uses a sentence to say something, the listener can come to know what the speaker says, by knowing the meaning of the sentence.

We can illustrate this with a simple example: Imagine two card-players who invent a system for communicating how good their hand is: Putting the hand to the forehead means few trumps, scratching ones beard means lots of trumps. The meaning of the signs is agreed on by both parties beforehand, and so when one signals, the other can understand the information in virtue of knowing the meaning of the sign. So for the sign-system to work, both parties must have agreed upon what counts as a sign, and on what the sign means.

The communication model invites us to think of natural language roughly along these lines: words and sentences are the signs used to communicate information from speaker to listener. Of course natural language is more complicated and speakers don't sit down to agree beforehand on the meaning of the signs, rather the signs have acquired their meaning over generations and somehow one picks up on the meaning by being part of the linguistic

community. The basic idea, however, attaching meanings which are known to the parties of the conversation to signs in order to use these signs in communication, is the fundamental picture behind more complicated models.

A well known picture, the so called semantic triangle illustrates the meaning relations between language, mind, and world, as conceived of in a communication model nicely:



On the picture suggested by the communication model and as pictured in the semantic triangle, linguistic items are associated with mental items, which, in turn, are about the world. The relationship between language and world is thus mediated by the mind.

Thinking of language as a tool for communication in this way gives rise to two philosophical questions and one linguistic question. The linguistic question is what the meanings of the particular signs, that is, word-types and sentence-types are and

how the meanings of the words and other subsentential components combine to constitute the meanings of sentences. The philosophical questions are more general. The first is: what are the meanings speakers of a language associate with a word or with sentences? The second question is: how do words and sentences acquire the meaning they have? The first question is a question of philosophical semantics, the second of metasemantics.[77] All three questions are probably not wholly independent of each other, for the answer to the linguistic question will constrain the answer to the question of philosophical semantics which will in turn inform the answer to the metasemantic question. Still, it is useful to keep the questions separate to avoid confusion.

Different answers to the question what the meanings of sentences and words are have been given. The natural answer, which we may take from the communication model and suggested by the semantic triangle is that sentences are associated with certain thoughts in the head of a speaker and of the listener, and that words are associated with the concepts of the speaker and the listener. What is so associated is the meaning of the sentence or word. However, to enable successful communication, what is associated with a sentence by the speaker should, in some sense, be the same as what is associated with the sentence by the listener. So what is associated must be a thought or concept with the same content. What these contents are, which speakers associate with words or sentences, is the matter of some philosophical debate.

---

[77] The distinction between semantics and metasemantics is fairly well entrenched. For an early distinction see Lewis (1975). For recent discussions see Burgess and Sherman (2014)

Some, taking a cue from formal semantics, think that propositions, the meanings of sentences, are functions from worlds to truth values, others, the proponents of Russellian propositions, conceive of propositions as sets of properties and objects, and Fregeans think that propositions, or thoughts, are constituted from the senses of the constituent words, where the sense of a word is not an object or a property, but a mode of presentation of an object.[78] The meanings of words, the concepts, are accounted for similarly, fitting the account of what propositions are. Who thinks of propositions as functions from worlds to truth values, will think of concepts as functions from worlds to extensions, and who thinks of propositions as sets of properties and objects, will think of the meanings of words as properties and objects, while the Fregean will think of the meaning of words as modes of presentation of objects.[79]

These different accounts have their strengths and weaknesses, but these are discussed by other philosophers at length elsewhere and the discussion need not be reiterated here.[80] The important point is that they all have a common starting point in thinking of words and sentences in a communication framework, that is, meanings are associated or assigned to words and sentences to enable these to communicate information.

---

[78] For an account of meanings as sets of worlds, see Lewis (1983) especially his "General Semantics", for a Russellian account see Salmon (1983) , and for a Fregean account see Evans (1982)

[79] These accounts of propositions are supposed to clarify the nature of propositions. In a more deflationary spirit one may also follow many of the arguments on which the debate turns, if one takes it to be about how we can best represent meanings in a model of our language. In that sense, meanings could merely model some of the aspects of the words of our natural language, as opposed to telling us what meanings are.

[80] A recent discussion of propositions by prominent philosophers in the field can be found in King, Soames, Speaks (2014)

On a simple picture, of which the toy language above is an example, knowing the meaning of a word or sentence involves associating the right concept or proposition with the word or sentence. So correctly associating a meaning with a word enables a listener of a language to understand the information the speaker attempts to transmit. This association or assignment of a word with a concept, or of a sentence with a proposition, is a contingent matter, and typically taken to be at least partly a matter of convention: a community of speakers decides on which word or sentence to use to communicate some piece of information, but could have used different ones instead.

This simple picture, which neatly associates word-types with concepts and sentence-types with propositions, however, is too simple to account for all intricacies of natural language, and thus needs to be extended. For in natural language, there is not just a simple and stable pairing between word-types and concepts, or sentence-types and propositions. The examples in the following should give the reader an idea of some complications a communication picture faces, and how it may be modified to account for these difficulties.

The most obvious examples where a simple matching between meaning and sentence, or meaning and word, fails, are sentences involving indexicals. Consider the indexical expression ‘I’ as it appears in the sentence “I am in London”. When I utter the sentence at the time of writing this, it means that I am currently in London, which is true. But if, say, my brother uttered the same sentence right now, it would be false, for he is currently in Germany. So the same sentence can transport two different pieces of information, that is, it can have different meanings. On the other hand it seems true to say that the sentence “I am in London”

does have the same meaning across different uses. But how can a sentence both have the same and a different meaning? The solution to this puzzle lies in distinguishing between two different kinds of meaning. The first is the meaning, sometimes called linguistic meaning, we associate with a certain word- or sentence-type of a language. This is the meaning of the word-type 'I', which we know as competent speakers. This meaning by itself, however, does not yet give us the information we wish to transmit, for in order to actually have such an informational content, we need contextual information, namely information on who the speaker of the sentence is when it is uttered on a particular occasion. Once context is filled in, we do have a full content. Thus we need to distinguish between two kinds of meaning: what speakers know, if they know the meaning, following Kaplan we may call this the character of a word- or sentence-type, and the information actually transmitted, which is the content of a word or sentence as used in a context.[81] We may think of the character of a word or sentence as a rule that tells the speaker which and how contextual information needs to be taken into account to arrive at the desired informational content of an utterance. In the case of 'I' this rule may be something like: The referent of 'I' is the speaker of the context.

It is quite controversial which words and accordingly sentences are context-dependent, and which contextual features should play a role. The most obvious contextual features are place, time, and speaker, but many more have been suggested. Contextualists

[81] The distinction as well as the terminology go back to Kaplan (1977). In a formal framework, one may think of the character of an expression as a function from contexts to contents. Where contents are again functions from worlds to extensions.

about knowledge, for example, have argued that the degree of justification required for something to truly count as knowledge is contextually determined by what is at stake in the situation in which the knowledge is gained or used.[82]

Whether a claim of context dependence is true in each individual case, is a matter of some controversy, but what should be well taken is that we need to distinguish between at least two kinds of meaning, the meaning stably associated by speakers of a linguistic community with a word- or sentence-type, that enables them to understand, taking contextual clues into account, what the informational content of an utterance is, and the informational content itself.[83] The character of a word or sentence is sometimes also called the linguistic meaning, for it is what can be stably associated with a word- or sentence-type of a language.[84] If a language does not contain any context-sensitive expressions, content and character do not come apart, but as soon as there are context-sensitive expressions, they do.

This slightly more complicated picture already considerably expands the simple communication picture introduced above. However, some natural language expressions still do not quite fit the expanded picture, because some expressions seem not to have a linguistic meaning at all, that is, there seems to be no meaning

---

[82] A nicely spelled out contextualist view of knowledge can be found in Lewis (1996).

[83] Gillian Russell (2008) suggests that we should not just distinguish between content and character, but also between content character and reference determiner, where the reference determiner is a meaning even more partial than character, which can be modeled as a function from a context of introduction to extensions. This kind of meaning is then supposed to handle cases of direct reference.

[84] Sometimes a further distinction is made between linguistic meaning and character, where the latter is narrowly defined as a function from context to propositions.

"in the head" to be associated with the word. An example of such expressions are proper names, for which it seems plausible that they refer directly, without the speaker or hearer having any stable or common rule in mind for picking out the referent. To use a name, a speaker does not have to know anything about its referent, she can just pick up on the use of the name in the community and make true assertions about the referent. One may, for example, know very little about Aristotle and not be able at all to pick him out with the information available, but still succeed in referring to him in conversation.[85] Examples like this suggest that not every expression of a language has a linguistic meaning, thought of as a set of conditions any competent speaker knows, but nonetheless the expression contributes to what a sentence says. So some expressions seem to contribute to what is said by a speaker without being associated with a content the speaker has in mind, but rather by being associated with the object it represents directly.

We have moved from the simple communication model, with its stable association of sentence-types with proposition, and words with concepts, to a distinction between two kinds of meaning, character and content, and on to a further peculiarity: expressions without apparent linguistic meaning. It seems safe to say that linguists as well as philosophers will find many more such peculiarities and further complications, as well as combinations of deference phenomena with context-sensitivity in natural language. For present purposes, these complications need not be further reiterated, they are discussed here to show that the

[85] A sketch of such a causal theory of reference can be found in Kripke (1980) and Putnam (1975). Other kinds of deference to the linguistic community, which not only involve proper names are described by Burge (1979b).

communication model has resources to deal with many complications, but also illustrate that despite the complications, the fundamental tenet of the communication model is the use of words as tools in successful communication. To facilitate communication, word- or sentence-types are contingently associated with meanings, even if the associated meaning is sometimes only partial.

The communication-model sketched here merits further exploration by linguists and philosophers alike. For the general study of representations, reference and its relation to truth, however, the communication model is, I will argue, not the most helpful model. It is here introduced as a foil to develop a quite different framework that abstracts away from the intricate question of how words, as tools in communication, succeed in transmitting information, and moves on to ask, how it is that representations generally relate to the world.

The reasons for moving from words to representations are methodological: If we want to study necessary truth and how it may arise from how reference is determined, it won't help to study sentences and their association with meanings only, for we want an account not of how some sentence-type with a certain linguistic meaning cannot fail to express something true, but of why what a particular sentence says in a context cannot fail to be true. The sentence 'I am here', for example, given its associated linguistic meaning, represents something true in any context in which it is uttered, but what it represents is something different on each occasion of use, and it is a contingent truth each time. So it is not the truth of a sentence or sentence-type which is of interest for an explanation of necessary truth, but the truth of what the sentence is used to represent.

Additionally, sentences are not the only kind of representations there are. Different kinds, for example mental representations, may represent the same thing as a sentence. So if a sentence represents something that is necessary, a mental representation that represents the same will be necessarily true as well. But if necessary truth was due to the connection between a sentence and its meaning, then this could not explain the necessity of the content of the mental representation. So, if there is to be any chance to explain necessary truth by some features of representations, the focus should not be on word- and sentence-types and their association with linguistic meanings, but on representations and their relation to what they represent quite generally.

2. Representations

Instead of taking word-types as the starting point of theorizing about reference, I suggest that taking token representations to be the primary objects of study is both more generally applicable, and more helpful for understanding the relationship between reference, truth, and necessity.

In what follows the theoretical framework in which representations are situated is outlined, and the relevant representational properties, reference and truth, discussed. To start, it will be helpful to get an intuitive grip on what representations are, by listing some properties of representations, and by giving examples.

Token representations have their representational properties essentially, and are of a type in virtue of having the same

representational properties.[86] Token representations can be either truth-evaluable or referential. Among the referential representations we can distinguish between singular and general referential representations. Singular representations essentially and rigidly refer to only one object, while general representations may refer to more than one object.

Truth-evaluable representations are constituted by referential representations, and the truth of the truth-evaluable representation depends on the (actual) reference of the referential representations. A truth-evaluable representation representing Tobi to be a lawyer, for example, is true just in case the referent of the referential representation ⌜Tobi⌝ is a referent of the referential representation ⌜lawyer⌝.[87]

This example already highlights a difference between classic semantic models and the present account of representations. On the standard picture introduced by Frege, predicates like 'is a lawyer' are not to be understood as referential, but as functions which yield a truth value relative to some input. The extension of such a predicate, or its so called reference, is then the class of objects for which it yields the truth value 'true'. Predicates (in so far as they are used as referential representations) in the current framework are understood as being general referential representations, which have the objects in their extension as referents. The truth of the truth-evaluable representation then depends on the relationship between the referents of constituent

[86] Which properties count as the relevant representational properties will be clarified in the subsequent discussion.

[87] In the following, I use corner-quotes (⌜…⌝) to talk about representations. Regular quotes will be used for quoting sentences ("…"), and single quotes ('…') for words or phrases.

referential representation. A defense of this stance will follow after having introduced the relevant terminology.

Representations are typically concrete objects, which have their representational properties essentially. A sentence as used on a particular occasion to make an assertion is an example of a token truth-evaluable representation, but the representation is not identical to that (token) sentence. Rather the sentence is a representation only in so far as it is used as a representation. Equally, some of the words of the sentence are examples of a token referential representation, but the referential representations are not identical to the words, which are referential representations only in so far as they are used as such.[88] A representational brain state is another example of a token representations, although the representation is not identical to the neuronal network realizing the representation.

So representations can have quite different underlying objects which realize them, but are not identical to any of them, as these underlying objects do not have the representational properties essentially. An analogy may help: The lump of clay that is formed like a statue is an example of a statue, but is not identical to any statue: a statue could be formed of quite different materials, it can be made of clay, it can be made of gold, etc. In this sense the clay-statue is only an example of a statue. But no statue is identical to the material of which it is made, for the material may have existed

[88] As mentioned before some may have been gesturing at a similar concept by talking about a 'thick conception' of words and sentences, which are supposed to be words and sentences cum meaning. For example Fine (1994), and García-Carpintero and Pérez Otero (2009).

before the statue came into existence, and may continue to exist after the statue is long gone.[89]

Even though different representations can have a quite different constitution, the examples used in the following will, due to the format of the current piece as a piece of writing, be examples involving words and sentences. When I talk about representations they will appear in corner quotes (⌜…⌝) and the word appearing in the corner quote should be understood as being a representation of the same kind as the one as which it is commonly used. I will in the following sometimes say that a representation is expressed by a word, phrase, or sentence. This way of speaking should not be confused with talk of sentences expressing propositions, prevalent in the communication-model.

That representations have their representational properties essentially means that their identity depends on them having certain representational properties, just like my being essentially a person means that my being me does not survive me ceasing to be a person, or, just like the statue, which, being essentially a statue, does not survive loosing the property of being a statue, even though its matter may survive.

That the representational properties are essential to the representation does not mean, however, that there is no explanation for why a particular representation has the representational properties it does. Just as there is an explanation for why the statue is a statue, and there is an explanation for why I

[89] There is of course a controversy about whether it is possible that more than one material object can overlap in spatiotemporal location, but I take this possibility to be well justified, and I will take it for granted in what follows. For a defense see Fine (2003), for a differing opinion see Almotahari (2014).

am a person, there is also an explanation for why a representation is the representation it is.

These remarks are a first starting point and should give the reader an idea of what kind of creatures representations are. Before moving on to the official definitions of the core concepts of the framework, consider the following example-pairs as further illustration of when truth-evaluable representations are token or type identical:

1:

a) The utterance “Snow is white”.

b) The utterance “Schnee ist weiß”.

2:

a) A sign which says “Free hotdogs here” above a hot-dog stand.

b) The same sign stacked away under the stairs.

3:

a) The utterance by me “I am here now”

b) The utterance by you “I am here now”

4:

a) The utterance “Hesperus is bright”

b) The utterance “Phosphorus is bright”

The first example illustrates a case of two sentences used as two distinct token truth-evaluable representations, which are of the same representation type. The utterance of “snow is white” uses different words from the utterance “Schnee ist weiß” and they are numerically distinct so they are distinct token representations. However, both assertions have the same representational properties and consequently they are of the same representation type.

The second pair illustrates a case where a token representation looses its representational properties, and thereby ceases to exist, even though the concrete object, the writing on the sign, continues to exist. The sign, while in use, indicates where free hotdogs can be obtained, but once it is stored away, it is not used any more as a representation and so it looses its representational properties and consequently ceases to be the representation it was.

In the third pair we have two distinct token representations, which use the same form of words with the same linguistic meaning to represent something different. Consequently, the two token representation, even though they have the same linguistic meaning and use the same form of words, are not of the same representation type.

Finally consider the fourth pair, which is another example of two distinct token representations that are of the same type. Their representational properties are the same, as both ⌜Hesperus⌝ and ⌜Phosphorus⌝ are singular representations of the same object, Venus, and so the two utterances are representations of the same type. The difference to the first case, in which different words were used to represent the same state of affairs, merely consist in the different words used to represent the same object being words of the same language. In the first example the words belonged to different languages.

The examples illustrate what it takes to be of the same representation type and what it takes to be a token representation. What should have become clear is that the starting point is different from the communication model in which words and sentences of a language and their associated meanings or propositions are the central objects of study. Although sameness

of representation type lines up nicely with what some philosophers have taken to be sameness of proposition expressed. This is as it should be, as propositions, at least in part, are supposed to be the full meanings of sentences. Representations are not propositions, however, as propositions belong to a different explanatory project. That this way of thinking about representations is more useful in the present context than a traditional communication-based framework utilizing linguistic meanings and propositions will emerge during the discussion, and will be argued for in detail in section six of this chapter.

In what follows, these first sketchy remarks are supplemented with a fuller theoretical framework that will be the basis for the account of modality to be developed. I will first focus on referential representations, before considering truth-evaluable representations.

3. Representations and Reference

The central relation that holds between a *referential* representation and what it represents is the reference relation. The reference relation holds between a referential representation and an object or a number of objects. A name token used on a particular occasion is an example of a singular representation, a representation that refers rigidly to one object, while general representations, for example predicates used as representations, typically refer to a number of objects. The singular representation ⌜Tobi⌝, for example, refers to the single individual Tobi, while the general representation ⌜blue⌝ refers to each blue thing.

That predicates, if used as representations, are said to refer to each object of their extension is uncommon, as they are more

standardly conceived of as functions from a domain to truth values. Their extension can then be defined as the class of objects for which the function returns the truth-value 'true'. While this construal of predicates as functions has technical advantages, it is by no means mandatory for an analysis of reference and its relationship to truth and may even be misleading about the right direction of explanation. It will emerge from the discussion that it is quite fruitful to conceive of both singular and general representations as referential and to explain truth as arising from the appropriate relationship between the referents, as opposed to taking truth as a semantic primitive which in turn elucidates reference.

Reference is a relation that holds between a referential representation and an object or objects, but it does not hold primitively. There is an explanation for why it holds: something constitutively explains the reference of representations. Thus the focus of the following discussion will be on reference determination. But before we can embark on this discussion, it is helpful to guard against a misunderstanding arising from two senses of reference determination, one epistemic, and one metaphysical. It is the purely metaphysical sense that is of interest to the present project.

In the epistemic sense of reference determination, reference determination is about a thinker having the epistemic means of picking out the referent of some word or phrase (in a context). So what determines reference in the epistemic sense is what a speaker or hearer knows that helps him or her pick out a referent of a word. In the simplest case a speaker knows a condition, sometimes supplemented by contextual information, that, if satisfied, picks out a referent or referents. This condition may

diverge from speaker to speaker and it may be that it picks out objects with the help of their contingent properties.[90] This epistemic sense of reference determination, however, is not the only one, and it is not the one salient for a framework making use of representations as opposed to words, for there is also a metaphysical sense of reference determination. What determines reference in the metaphysical sense, are the properties constitutive of the reference relation holding between referential representations and the objects in their extension.

Whenever a reference relation holds between a representation and some object or objects, it holds in virtue of some properties of the representation, call these *the reference determining properties* and in virtue of some properties of the referent(s), call these properties of the referent(s) the *reference realizing properties*. The reference determining properties and the reference realizing properties together are constitutive of the reference relation between a representation and an object or objects.

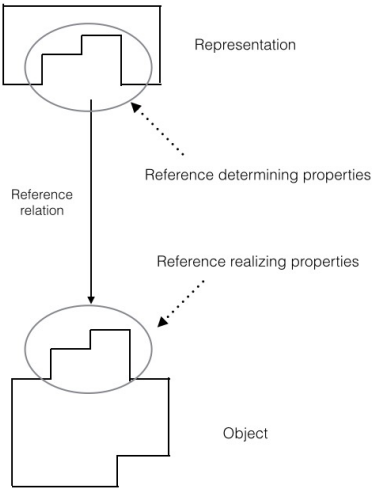
While the reference determining properties and the reference realizing properties work together to establish a reference relation, it is the reference determining properties that determine which properties the referent needs to have for the reference relation to be established. So the reference determining properties are explanatorily prior to the reference realizing properties in that they determine which realizing properties the referent must have for a reference relation to be established. Note, however, that the reference determining properties do not determine the referents to have the reference realizing properties. They are had by the objects prior to anyone referring to them. It is just the fact that

---

[90] Burge (1979a) may be read as arguing that we should think of Fregean senses along these epistemic lines.

they are referred to by a representation which makes these particular properties of an object the reference realizing properties vis à vis a representation.

As an analogy, consider the relation of fit as it holds between a lock and a key: For the relation of fit to be established between a lock and a key, the lock as well as the key need to have the right form. The properties of the lock will determine what properties the key must have to fit into the lock.[91] We may say that the properties of the lock that contribute to establish the relation of fit are the fit-determining properties, while the properties of the key that contribute to the relation being established are the fit-realizing properties. Consider the following picture:



[91] Also, the form of the key determines what a lock must look like into which it fits. This direction of determination, however, corresponds to a different explanation. Equally in the case of referent and representation.

The picture illustrates the reference relation as it holds between a representation and an object. The bulge on the object side represents the reference realizing properties while the notch on the representation side represents the reference determining properties. Whenever the bulge fits into the notch, so to speak, the reference relation is established.

It may happen that there is no object that, so to speak, fits into the notch of a representation. In this case no reference relation is established, but the representation still has reference determining properties that determine what properties an object would have to have to be the referent of the representation.

The important feature of the reference relation that is illustrated by the analogy to a relation of fit is that both relations hold in virtue of properties of their relata, and that the properties of the one relatum determine the properties the other relatum must have for the relation to be established.

Analogies, however, can be treacherous, as they may misrepresent some features of the analogon. The function of the present analogy is to illustrate how a relation can be established in virtue of properties of the relata, and how properties of the one relatum can determine what properties the other relatum needs to have for a relation to be established, it is not the claim that the reference relation is a relation of fit, and it is not the claim that reference is somehow established by relations of fit between a representation and its referent(s). Such suggestions, reminiscent of a picture theory of meaning, would be too narrow and therefore implausible.

Type-identity for referential representations

One important question that can be tackled with the developed concepts, is a criterion of type-identity for token representations. The focus is presently on referential representations and the criterion of type identity will be extended to truth-evaluable representations in the next section.

Intuitively we would like to say that referential representations are of the same representation type, if they represent the same object(s). However, we cannot just define sameness of type of a representation in terms of sameness of object(s) that it refers to, as there are representations that happen to have the same extension, but represent something different.[92] Also, defining sameness of representation type in terms of sameness of reference determining properties is not helpful, because it seems that different representation tokens may have different reference determining properties and nonetheless represent the same object(s).

We can, however, look to the reference realizing properties for a criterion of sameness of representation type. Since the reference realizing properties are those properties of an object that are responsible for its being the object referred to by a given representation, it seems reasonable to expect that representations whose reference determining properties determine the same reference realizing properties are guaranteed to refer to all and only the same objects. In turn, it seems to be the central requirement on representations of the same type that they are guaranteed to refer to the same things in virtue of how their

[92] The classic example is Quine's of 'creature with a kidney' and 'creature with a heart' (Quine 1986). As it happens, everything that has a kidney has a heart, and so both have the same extension, but they obviously represent something different.

reference is determined. Consequently, sameness of reference realizing properties is a good candidate for defining type-indentity of referential representations.

> *Type-Identity of referential representations:* Two token referential representations are of the same type, iff their reference determining properties determine the same reference realizing properties.[93]

This criterion of type identity is a criterion for singular representations as well as for general representations. Before moving on to clarify the relationship between reference determining and reference realizing properties, and giving some examples of what these properties may be in specific cases, lets take stock.

Token referential representations stand, except in cases of reference failure, in reference relations to the objects they represent. Some representations refer only to one thing, others refer to multiple things. In both cases the reference relation, if it holds at all, holds between the representation and the referent(s) in virtue of properties of the representation and in virtue of properties of the referent(s). The former are the reference determining properties, the latter the reference realizing properties. The reference determining properties are explanatorily prior in that they determine what the reference realizing properties

[93] We may think of sameness of representation-type as sameness of meaning. But we shouldn't confuse this notion of meaning with notions of meaning as they figure in theories of meaning based on the communication model, or theories in which meaning has strong epistemic import.

are, if a reference relation holds, or would be, if it would hold.[94] Sameness of representation-type can then be defined in terms of sameness of reference realizing properties. For convenience I will in the following talk about referential representations determining reference realizing properties or about reference realizing properties determined by the representation. Properly expanded, this should be understood as saying that the reference determining properties of the representation determine what the reference realizing properties would be, if a reference relation were to hold.

4. From Reference Determining to Reference Realizing Properties

Quite a bit has been said about the new concepts, but I have not been very specific about what reference determining properties, and even more importantly, what reference realizing properties are in specific cases as well as how the reference determining properties succeed in determining what the reference realizing properties of a referent are.

Before clarifying this with some examples, I argue that some top down considerations already reveal quite a bit about what the reference realizing properties of objects may be. Following these general top-down considerations, I take a look at singular and general representations and outline a way of arriving at reference realizing properties for these kinds of representations. To do this, a bottom up approach is used, utilizing two prominent kinds of philosophical theories of reference, the description theory, and the causal theory.[95] Both theories are adapted to the present purpose,

---

[94] One way to put this is that the reference relation has a direction from the representation to the object, which is mirrored in the direction of determination.

[95] The classic version of a description theory for names is Russell's (1905). For a refinement in terms of clusters of descriptions see Searle (1983). Outlines of the causal theory can be found in Kripke (1980), Putnam (1973), (1975).

for they are more naturally situated in a communication-framework. Both are used here for purposes of illustration only, and I do not wish to advance any thesis about which theory may be correct for a particular representation. Both theories have quite serious problems, and the simple sketches used in the following should at no point be taken to provide the whole story about reference determination.[96] In principle, the present framework can stay completely neutral on which theory of reference is correct, as long as some reference realizing properties are determined by the reference determining properties.

Determining reference realizing properties: top down

The reference realizing and the reference determining properties of the representation and the referent constitutively explain the holding of the reference relation between the two, provided a reference relation holds at all. This characterization already allows to partially delimit what the reference realizing properties of an object may be:

An object is the referent of a singular representation, if and only if the object represented by the singular representation exists, and an object is the referent of a general representation, if and only if the object has the feature represented by the general representation.

Since the reference realizing properties are the properties of an object which are constitutive of reference, the reference realizing properties are had by any object, if and only if it is the referent of the relevant representations. And so if and only if the referent of a referential representation exists, does anything have the relevant reference realizing properties. This has some implications for

[96] For a discussion of some of the problems of causal theories of reference see Evans (1973),(1982).

what the reference realizing properties can be. Let us first consider singular representations. Assuming that most objects that are the referents of singular representations have a temporal extension, it follows that an object has its reference realizing properties during its existence.[97] Thus the reference realizing properties of singular representations are properties the referent has during its existence. Once the referent looses the properties, it ceases to be the referent of the representation, and since if something is the referent of a singular representation, it is so during its lifetime, it ceases to exist, if it looses the reference realizing properties. Plausibly this finding from the temporal case also holds in the modal case. The reference realizing properties of an object are had by it across counterfactual situations. So the reference realizing properties are not merely among the properties an object has during its temporal existence, but they are among the properties an object has 'during' its modal existence, that is, it necessarily has the reference realizing properties, if it exists.[98]

This already cuts down on the candidates for what the reference realizing properties for singular representations may be proceeding from nothing but the general characterization of

---

[97] The temporal reading suggested here is a little tricky, as reference relations arguably hold across time (unless you are a presentist): we can refer to future objects, and, even more uncontroversially, we can refer to past objects which no longer exist. So there seems to be no clear time at which the reference relation holds or doesn't hold. But to get to the relevant temporal reading, we can focus on the referent, the object which has the reference realizing properties, and say of it, that it is a referent of a singular representation only during the times at which it has its reference realizing properties.

[98] Note that this need not hold for each reference realizing property singly, for it may be that the set of necessary and sufficient reference realizing properties determined by a representation can only be characterized disjunctively. Thus, when I talk about the reference realizing properties of an object, this should be understood as talk about the necessary and sufficient properties constitutive of reference.

reference realizing properties and general assumptions about the reference relation. However, this account cannot and should not be understood as a reduction of reference realizing properties to necessary properties of an object, for there are a host of necessary properties an object has which are clearly not reference realizing. So being a necessary property is not a sufficient condition for being a reference realizing property of a singular representation, however, it is a necessary condition on being the reference realizing properties of an object that the represented object has them of necessity, if it exists.

Note that the explanatory relations do not run from being necessary to being reference realizing. To the contrary, the direction of explanation runs the other way around. That a property is a necessary property does not determine that it is reference realizing, it is merely an epistemic criterion for finding out what the reference realizing properties may be, provided we know what the necessary properties are.

These findings for singular representations can be partially transferred to general representations. The difference is that while the referent of a singular representation is the referent during its existence, the referent of a general representation only has the relevant reference realizing properties while it is a referent of the general representation, that is, while it has the feature represented by the general representation. Still, while an object is the referent of a general representation, the reference realizing properties are among those properties the object has only during that time. The same holds for the modal dimension: If it would be the referent of the representation, it would have the relevant reference realizing properties.

These general considerations somewhat delimit the candidates for what the reference realizing properties of an object may be and point to a certain similarity between essential properties and reference realizing properties that will be picked up again in chapter eight. It is not enough, however, to delimit what the reference realizing properties determined by a particular representation may be. To get closer to an answer, a more detailed account using a bottom-up approach is provided in what follows.

A more detailed account: singular representations

Singular representations essentially refer to only one object throughout the object's existence. Thus the reference realizing properties must be properties that distinguish the object from all other objects, and, as we have seen above, they must be among

the necessary properties the object has. These requirements lead to two challenges.[99]

The first challenge is that the purely qualitative properties of an object typically cannot serve to distinguish it from all other objects, and even if they actually can, they are usually not properties the object has necessarily or during its entire existence. To get this problem into focus, let us look at two pieces of paper which come from the same production line, and let the first be

---

[99] Standard examples of singular representations are names. However, definite descriptions like "the chancellor of Germany" are also used as singular representations. Definite descriptions are interesting, because they seem to refer to an object in virtue of an object satisfying a description that typically mentions contingent features of the object. Thus, it may seem that the reference realizing properties in this case are properties the object has merely contingently. This impression is false, however, for there are at least two readings of definite descriptions: a rigidified reading and a non-rigidified reading. On the rigidified reading, the description "the chancellor of Germany" refers to Angela Merkel. If it is to be the case that the chancellor of Germany is identical to Angela Merkel, then the two representations have to be of the same type, and thus have to determine the same reference realizing properties. So on the rigidified reading, the reference realizing properties determined by "the chancellor of Germany" do not include the contingent feature of being the chancellor of Germany, rather they are identical to the reference realizing properties determined by "Angela Merkel". There is a non-rigidified reading, however, on which the reference of "Angela Merkel" and "the chancellor of Germany" come apart. This reading is exploited by Kripke (1980) in his modal and temporal counterexamples to the description theory of names. On this reading, the reference realizing properties determined by "the chancellor of Germany" really are being the chancellor of Germany. But since "the chancellor of Germany", on this reading, refers to different people across time and modal space, "the chancellor of Germany" does not refer to Angela Merkel, rather whoever happens to be the chancellor of Germany.
So we have to be careful which reading we take to be salient. Typically, I suppose, the rigidified reading is intended. But the non-rigid reading is certainly also possible, for it seems salient when considering some counterfactual situations. The trouble with the description theory of names was that it combined an intuitively rigid reading with a theory that yields the wrong reference realizing properties for this reading. In either case, we must take care to properly distinguish the different readings when analyzing what the reference realizing properties of such representations are.

represented by the singular representation ⌜A⌝, and the second by ⌜B⌝. Let's assume the two pieces of paper are qualitatively identical, so no properties like size, weight, surface, etc. distinguish them. How is it that the representation ⌜A⌝ refers to the one piece of paper and ⌜B⌝ to the second? Initially one may think that there is an easy answer to this question, for did I not say that the first piece of paper is the referent of ⌜A⌝, and the second the referent of ⌜B⌝? So what distinguishes them is that A was the first piece to come out of the machine, and B the second. However, while this does indeed distinguish the two pieces, the property of being the first piece of paper is not a property A has necessarily, for it could have been that A came out second, and B first. So this feature of the piece of paper cannot be what is constitutive of reference.[100]

The second challenge is to find some feature or features of the object which explain the holding of the reference relation during the object's existence in temporal space, as well as in modal space. That is, they are properties the referent cannot loose, if it is to remain the referent of the representation, and since singular representations refer to objects if and only if they exist, they are properties the object cannot loose without ceasing to exist.

These two challenges for finding the correct reference realizing properties for singular representations, distinguishing the object

---

[100] Philosophers who are sympathetic to the essentiality of origin may take this to be false and instead believe that the origin of the piece of paper is indeed what distinguishes the two. A thesis of the essentiality of origin, however, is not necessary for making the relevant distinctions, or so I will argue, and consequently I will not commit myself to this controversial thesis, even if it may offer a quick solution to some troubles.

from every other object there is and 'tracking' the object's temporal and modal extension, are discussed in what follows with the help of an example.

Consider the name 'Tobi' which I will in the following take to be used as a singular representation for Tobi. As such the representation token refers to Tobi, if and only if Tobi exists. I will assume a simple sketch of a causal theory of reference along the lines suggested by Kripke on which there is a chain of uses of the name 'Tobi' as a representation of Tobi up to the use here, where each token representations inherits its representational properties from the use before. The chain goes back to an initial baptism at which the reference of the name was first fixed to Tobi. I do not take this sketch to be an accurate account of how reference is to be explained for all singular representations, however.[101] There are many more options, and the inheritance-of-representational-properties-account suggested here is almost certainly too simple.[102] But this is not our primary concern presently and must be left to empirical research in each specific case. Considering the simplified model case will nonetheless be instructive.

Representations in the chain of uses inherit their reference from the use before, that is, in virtue of standing in a certain causal connection to token representations before, they are of the same type. This part of the account is not particularly interesting for finding out what the reference realizing properties may be: They

---

[101] One further option may be rigidified descriptions.

[102] In some cases inheritance may fail, or the representational properties may change during the use of a name-type across time, as the famous Madagascar example by Evans (1973) shows. So it may happen that during a chain of uses, the uses of the same name types are of a different representation type.

will be the same as those of the representation used before. This is all the causal part of the causal theory is able to tell us.

A more interesting part of the causal theory for finding out about the reference realizing properties is the initial baptism at which the reference is fixed to the object. Most accounts of the causal theory are not very explicit about the baptism, for it seems that it is quite easy to baptize something. Kripke, for example, observes that we do not have to do much to succeed in giving a name to, say, a person.[103] When asked what they want to call their child, parents just have to say a name and henceforth the name is used to refer to the child.

However, just because we do not have to do much to endow a name with a referent, does not mean that only a few things must be the case for a baptism to succeed, for somehow it needs to be determined what the representation is supposed to be a representation of. And since many objects are present at the baptism, there must be something that fixes the reference to one of these objects. That is, something which determines the reference realizing properties such that the appropriate object is the referent of the representation. Merely pointing to something at the baptism doesn't help, for we can point as hard as we want and not distinguish, say, a mere time slice of a person from a person.[104]

---

[103] Kripke (1980) p.93ff.

[104] A version of this problem is by Devitt and Sterelny called the "qua-problem". For a discussion of the qua-problem see Devitt and Sterelny (1987), Stanford and Kitcher (2000), and most recently Thomasson (2015). For a slightly different approach to the topic see (Wiggins 2012). Sometimes the problem is used to promote some version of the description theory, or the weaker claim that there must be some descriptive element to the meaning of directly referential expressions. This is not the aim of the present discussion, since our focus is on representations and how their reference is fixed, not in what linguistic meaning words may have.

One option for determining the reference realizing properties may be to use what Kripke calls a reference fixing description. These descriptions are used at an initial baptism to fix the reference at the baptism. The problem with this, however, is twofold.

First, and most importantly, the description used to fix the reference typically uses contingent features of the object the reference is fixed to. Consequently the features listed in the reference fixing description, if contingent, cannot be the reference realizing properties of the object.

Second, the description doesn't even have to apply to the object that is baptized to be used as a reference fixing description: Consider the baptism of a mountain by saying "That volcano over there shall be called 'Mt. Tobi'". As it turns out, there is no volcano there, it is just a peculiarly shaped regular mountain. Still the baptism may succeed and henceforth the words 'Mt. Tobi' are a representation that refers rigidly to the indicated object, that is, the mountain. So the properties used in a reference fixing description are typically not the reference realizing properties.

Nonetheless there must be something in the story about how a representation gets its referent that determines the reference realizing properties. Luckily, there are a few more things that have to be in place for a baptism to succeed that have not been mentioned and which give us a better clue as to what the reference realizing properties may be. These facts can be identified by considering what the representational system is used for and what singular representations are supposed to represent in that system. This is most explicit, if we consider representational systems that are build for a specific purpose. Consider for example a meteorological model which predicts amounts of rain, wind, clouds, etc. In the model, there are representations for

clouds, rain, wind, and their respective amounts. Part of the reason the representations in that model do represent these things is that the model is build for the purpose of modeling these features of the weather.

Not all representational systems are purpose-built. Language as well as our brains, considered as representational systems, are almost certainly not built by anyone for any explicit purpose, rather they are, we may assume, the product of evolution. This does not mean, however, that these systems are not used for some purpose which in part determines what the objects represented in that system are.[105] We use language, for example, to coordinate our social interactions, and we use our brain for navigating the world of our experience. This gives us some clues as to what the referents of our singular representations may be, what the objects we seek to track with singular representations may be.

Turning back to our example of names, a few uses of these singular representations come to mind: We use names for persons, we use them for animals, we use them for places, we use them for other kinds of objects of our life-world. In the present example the name 'Tobi' is apparently a representation of a person, and so this gives us a clue as to what a reference realizing property, necessary for reference of the singular representation ⌜Tobi⌝ may be: being a person. A reference relation holds between ⌜Tobi⌝ and Tobi, only if Tobi is a person. Once he ceases to be a person, the

---

[105] This basic insight is part of the foundation of the broad research project of naturalized semantics, and more specifically teleological semantics. I do, however, have no intention of endorsing this entire approach to semantics, and its quite reductive ambitions. Rather I stay neutral on what will in the end be the right way to account for the reference of our representations. For more on naturalized semantics see Loewer (1997), for a famous teleological theory in terms of evolutionary fitness see Millikan (1989).

name refers no longer, and it did not refer before Tobi was a person.

Note that I am not claiming that the actual names we use must always refer to persons. We do sometimes use names differently, when we, for example, use the name a person had during his or her life to refer to his or her body after the person has died, and one may be inclined to say that the referent of a name is the same on both uses. So the present suggestion should be taken as a falsifiable empirical hypothesis about what we use our singular representations for. Nonetheless, the hypothesis that some names, in particular the name 'Tobi' in the example, refer to a person, and only to a person, does not seem to be too far from the truth, and it will be a working assumption in what follows.

Having identified this property of Tobi helps to solve the second problem of finding something that tracks Tobi's extension in temporal and modal space. However, while having the reference realizing property of being a person is a necessary condition for being the referent of the singular representation ⌜Tobi⌝, it is not sufficient, for there are many persons, and the representation refers to only one of them. So what property or properties of Tobi make him the unique referent of ⌜Tobi⌝? This question does not have an easy answer, for as was pointed out above in the paper-example, no conjunction of purely qualitative properties seems up to the task of individuating any object uniquely.

Here the baptism-picture may help: Tobi is the very person which is actually present at the baptism, the one we give the name to, no other person was given the name at this very baptism. Once the reference is thus fixed, the singular representation refers to that very person.

So the property which succeeds in picking out the person uniquely is that it is the very person *actually* present at the baptism. So the reference realizing properties for being the referent of ⌜Tobi⌝ may be characterized as being the same person which was actually present at the initial baptism, or, for short, the person identical to Tobi.

This characterization is rather unilluminating as long as no more has been said about what it takes to be identical to Tobi, or more generally, what it takes to be the same person or object actually present at the baptism. However, a few accounts from the literature come to mind, which may help to account for the identity. First, there is the thesis of the essentiality of origin, which says that in order to be the same object, it must have the same origin.[106] One could, so the intuition goes, not have been born of different parents and still have been the same person. Other accounts, mainly interested in what it takes to be the same person or object over time, account for identity with some kind of continuity, either of the body, or of the mind. On those accounts, what it takes for a person (or object) at one time to be identical to itself at another time, is to have some continued bodily or mental integrity. How exactly this is to be spelled out, is a further difficult question and the matter of some debate; a debate I do not intend to enter into presently.[107] A third view may take a cue from essentialist views of natural kinds, and take some feature of the

---

[106] Since Kripke (1980) reintroduced the thesis of the essentiality of origin, it has received quite a bit of discussion. For a defense see Salmon (1981). For a more critical view, see Mackie (2006).

[107] For an overview over the views and debates on personal identity, see Olson (2016)

person's internal constitution, for example the person's DNA, to be what is constitutive of being the same person.

There is no need to take sides as to what ultimately accounts for the identity of objects and there may be quite diverse answers for different kinds of objects. In either case, however, there seems to be something philosophically interesting to say about what constitutes the identity of objects of the kind in question. So I will let the reader pick his or her favorite view for each particular case, but assume that there is something more to be said about what it is that makes an object the very object it is, apart from the brute statement of identity, which can then be utilized to characterize the reference realizing properties.

Let me take stock: the reference determining properties of singular representations must suffice to determine the reference realizing properties of the referent such that the object is the unique referent throughout its existence. What determines reference for singular representations are both what is explicitly appealed to in theories of reference, for example, the standing in some appropriate relation to other representations, but also facts about what the representational system is used for. In the case of names in our language I suggested that the determined reference realizing properties in many cases include the property of being a person, and in other cases the natural kind the referent falls under. This property tracks the temporal and modal extension of the referent and is thus a good candidate for being among the reference realizing properties. In addition it is plausibly determined by the reference determining properties, if we take general facts about the representational system into account. To make sure that the representation is singular, that it refers to only one object, a connection to the existing object needs to be

established, it must be the object which is actually present at the baptism.

A more detailed account: general referential representations

General referential representations potentially refer to a number of objects. This makes accounting for general representations somewhat easier, because uniqueness does not have to be guaranteed. The representation ⌜red⌝ refers to each and all red things, the representation ⌜ball⌝ to each and every ball, and similarly for other general representations. The referents are the referents of these representation in virtue of having the appropriate reference realizing properties. Since objects can be the referents of general representations at some time during their existence, but not at others, the reference realizing properties are typically not properties the referents have necessarily. However, they are properties an object has while it is the referent of the representation and only while it is the referent of the representation.

So what are the properties the referents of general representations must have to be the referents of the representation? Initially one may take the answer to be quite trivial: To be the referent of ⌜red⌝, an object needs to be red. To be the referent of ⌜ball⌝, an objects needs to be a ball. And in order to be the referent of ⌜red ball⌝, an object needs to be red and a ball. This trivial story usually does work, but it is interesting to see why it works, why it might not be so trivial after all, and why we can often say something non-trivial about what is constitutive of reference.

The reason the seemingly trivial story works, is because in our language the use of predicates and the use of words for properties

is systematically related: If some predicate used as a general referential representation refers to some object, we can take the word and turn it into the name of a property: the property unifying the extension of the predicate. This also works the other way around: a name for a property or feature of an object can be turned into a predicate and thereby be used to refer to the objects which have this feature. This relationship between the words used to represent properties and words used to refer generally is quite useful for a language to have, but it also leads to the apparent triviality of the characterization of the reference realizing properties. For to characterize the reference realizing properties, we have to use language once more.

The resulting triviality of the characterization, however, is not indicative of there not being an explanatory relation between the property of being red and being a referent of the general representation ⌜red⌝, it is merely indicative of a special feature of the language we use to characterize this explanatory relation.

Still, there are many cases in which we can do better than to give a trivial story, for we can give a fuller account of what it takes to be the referent of a general representation. Lets look at two examples where non-trivial reference realizing properties for general representations can be found. The first utilizes a description theory of reference, the second uses the sketch of a causal theory of reference.

Lets first take a look at an example to which a description theory may apply: the representation expressed by the word ‘vixen’. On the description theory, the word used as the representation ⌜vixen⌝ has its representational properties in virtue of being appropriately associated with certain mental representations. In

the present case, let us assume, it inherits its representational properties by being associated with the mental representations ⌜female⌝ and ⌜fox⌝ such that it refers to the objects to which both ⌜female⌝ and ⌜fox⌝ refer.

The description theory gives an account of how one representation inherits its representational properties from others: By being appropriately associated with other representations, it inherits their representational properties such that the reference realizing properties of the representation are the conjunction of the reference realizing properties of the associated representations. While this theory does not tell us directly what the reference realizing properties of the referents are, it moves us one step closer to an answer. For if the referents of ⌜vixen⌝ are supposed to be referents of both ⌜female⌝ and ⌜fox⌝, it must be the case, that the referents of vixen have both the reference realizing properties of ⌜female⌝ and ⌜fox⌝. Consequently, we arrive at some non-trivial reference realizing properties: the referents of ⌜vixen⌝ have the reference realizing properties being female and being foxes.

Another theory of reference that can also be utilized for general representations is a causal theory of reference. It proceeds along similar lines as the causal theory for singular representations. Again the representational properties of one representation are inherited from the representations used before to which the representation is connected by some chain of uses. The chain of uses terminates at an initial baptism where a sample or example is

first indicated as a referent of the general representation.[108] The stock example for this kind of theory is the general representation ⌜water⌝, and so I will use it here.[109]

In the literature, it is not always clear whether the general representation ⌜water⌝, which has as referent everything that is water, or whether the singular representation ⌜Water⌝ referring to the kind water, or the property of being water is at issue. Presently, as in Putnam's original discussion, I take it that the theory provides us with an account of how the reference of the general representation is established.

So here is the simplified story of how the general representation ⌜water⌝ gets its representational properties: The representation stands in a chain of uses of other token representations of water, going back to an initial baptism at which a sample of water was baptized thus: "Samples of the same natural kind as this are called 'water'!". After that baptism, uses of the term 'water' are general representations of the objects which are of the same natural kind as the indicated sample. So what it takes to be the referent of the representation ⌜water⌝ is to be an object of the same natural kind as the one actually indicated at the baptism. But we can be even more specific. For what it takes to be of the same natural kind as the water indicated at the baptism, is to have the same molecular

---

[108] Once more the simple theory is most likely false for most representations we actually use. It is nonetheless instructive.

[109] That 'water', as a natural kind term, should be treated in this way was famously argued for by Putnam (1975).

structure, or so common orthodoxy goes.[110] Thus, what it takes to be water is to have the molecular structure H_2O. So the reference realizing property of objects to which the general representation ⌜water⌝ refers is having the molecular structure H_2O.

In the example it is assumed that the representation ⌜water⌝ has as referent the members of a natural kind which are of a natural kind in virtue of having the same molecular structure. This may be true for ⌜water⌝, but this may not be the case for every representation whose reference is determined in a similar way. Which underlying properties turn out to be the reference realizing properties of these representations, will depend, as in the case of singular representations, on a number of facts about the representational system in which they appear as representations.

This concludes the present discussion of referential representations and the relationship between reference realizing and reference determining properties. Some more suggestions about what reference realizing properties in specific cases may be will be made along the way, especially when discussing examples of necessary truths in chapter six and in the discussion on essentialism in chapter eight.

5. Representations and Truth

There are referential representations, and there are truth-evaluable representations. Token referential representations come together to form token truth-evaluable representations. The truth of these

---

[110] Even if it may be true for the present example that kind-membership is due to molecular, or otherwise internal structure, it may be that kind-membership is not generally so determined. I do not take a stand on the issue here. In principle the causal theory seems compatible with many different ways for individuating kinds. For a more complex approach see Stanford and Kitcher (2000).

truth-evaluable representations depends on whether the relationship between the referents of the constituent representations is as the representation represents it as being.

Suppose, for example, that a truth-evaluable representation is constituted of the referential representations ⌜raven⌝ and ⌜black⌝ in such a way that it is required for the truth of the truth-evaluable representation that the referents of ⌜raven⌝ are also referents of ⌜black⌝. This truth-evaluable representation is true in virtue of the appropriate relationship between the referents of the referential representations. In particular, the truth-evaluable representation is true, just in case all the referents of ⌜raven⌝ are referents of ⌜black⌝, that is, if all ravens really are black. These requirements on the referents of the constituent referential representations are the *requirements for truth* of a truth-evaluable representation.

Referential representations can constitute truth-evaluable representation in different ways, thereby giving rise to different requirements for truth. The two referential representations ⌜raven⌝ and ⌜black⌝ may, apart from the first possibility above, also constitute a truth-evaluable representation such that it is required for its truth that the referents of ⌜black⌝ are all referents of ⌜raven⌝ which would be true, if everything black was a raven. Different ways of combining the two referential representations from the example into truth-evaluable representations may give rise to still other relationships between the referents of the representations required for the truth. So it may be that an appropriately constituted representation requires for its truth that

most referents of ⌜raven⌝ are referents of ⌜black⌝. Many more ways of constituting a truth-evaluable representation from referential representations come to mind when considering these examples. What unites these ways is that they all require some appropriate relationship between the referents of the constituent referential representations for their truth.

How truth-evaluable representations are constituted from referential representations and how this constitution gives rise to different requirements for truth, depends on the kind of representation at issue. In the case of representations in language, the requirements for truth are indicated, for example, by the copula ‘is’, by the order in which the referential representations appear, by logical vocabulary, and by quantifiers.

The assumption that quantifiers together with structural properties of sentences and logical vocabulary can be, and are used to indicate the relationship between the referents required for the truth of a representation is consistent with the treatment of quantifiers by linguists and philosophers alike: Quantifiers in subject-position are to be understood as denoting relations between those sets which are the extension of the words of a sentence: the universal quantifier denotes the subset relation, the existential quantifier the relation of non-disjointness.[111] Other relations between sets are denoted by other quantifiers. For example the sentence, “Every raven is black”, is true, iff the extension of ‘raven’ is a subset of the extension of ‘black’. And

[111] See Heim and Kratzer (1998), p. 147ff. They attribute this relational view of quantifiers to Aristotle, and trace it through Leibniz and Frege. Heim and Kratzer also argue at length that quantifiers cannot be understood as denoting either individuals or sets (p. 131ff.).

"There is a beer in the fridge." is true, iff the set of beers is non-disjoint from the set of things in the fridge.

So in the case of sentences that are used as representations, it seems to be well established that quantifiers are used to indicate the relation the referents of the referential representations need to stand in, in order to be true. How this is done in the case of other kinds of truth-evaluable representations, for example thoughts, is a further question, which will need to be answered by looking at other kinds of representations.

The important point presently is that referential representations come together in various ways to form truth-evaluable representations. Depending on how they come together, different relationships between their referents are required for the truth of the representations. In each case, the truth of the truth-evaluable representation will depend on the referents of the referential representation constituting it. The relationship between the referents of the constituents required for the truth of the representation are the requirements for truth of the representation.

With this concept of requirements for truth in hand, we can go on to define sameness of type for truth-evaluable representations just as we did for referential representations:

> *Type-Identity of truth-evaluable representations:* Truth-evaluable representations are of the same type, iff they are constituted by referential representations of the same type, and the same relationship between the referents of these representations is required for their truth.

This definition is the natural extension of the definition of sameness of type for referential representations.

The present discussion provides the basic concepts for an explanation of necessary truth. Before moving on, some remarks on complex truth-evaluable representations and how they may be broken down into their constituents are made. To conclude the introduction of the new concepts, I assess the utility of the developed concepts.

Complex truth-evaluable representations

The examples discussed above all had a very simple structure, but truth-evaluable representations may be quite complex. Examples illustrating this are representations that are expressed by sentences involving logical connectives. Let us start with a comparatively simple example of a representation that illustrates this complexity, the representation expressed by the sentence "If Peter is grey-haired, then Peter is old". This truth-evaluable representation apparently consist of the singular representation ⌜Peter⌝, another singular representation of Peter of the same type, the general representations ⌜grey-haired⌝ and ⌜old⌝. In addition there is the logical connective if…then, whose role in the sentence used to express the representation needs to be clarified.

What makes the example initially puzzling, is that the representations ⌜Peter⌝ and ⌜old⌝, as well as ⌜Peter⌝ and ⌜grey-haired⌝ apparently already form two truth-evaluable representations, which are in turn connected by the logical connective to constitute another truth-evaluable representation. Thus it may seem that truth-evaluable representations are not always constituted by referential representations, but may also be constituted by other truth-evaluable representations. This, however, is only partially right: all truth-evaluable representations

should be thought of as primarily constituted by referential representations, even though a proper subset of the representations which partly constitute a truth-evaluable representation could sometimes constitute truth-evaluable representations themselves.

The reason we should always conceive of complex truth-evaluable representation primarily as constituted of referential representations is that the truth of the truth-evaluable representation ultimately depends on the relationship between the referents of the constituent referential representations. Consequently, the logical connective should be read as telling us something about the requirements for truth of the complex representation, which are given in terms of the referential constituents.

If this is right, the requirements for truth in terms of the referents of the referential representations are the following: ⌜If Peter is grey-haired, then Peter is old⌝ is true, iff the referent of ⌜Peter⌝ is a referent of ⌜grey-haired⌝ and the referent of ⌜Peter⌝ is a referent of ⌜old⌝, or the referent of ⌜Peter⌝ is not a referent of ⌜grey-haired⌝.

The example demonstrates that the complexity of a referential representation gives rise to complex requirements for truth. One may, however, be inclined to believe that the sentence above should be understood to express a different representation; a representation which represents the logical relation 'if…then' as holding between the two truth-evaluable representations ⌜Peter is grey-haired⌝ and ⌜Peter is old⌝. For this account of the representation to work, we would have to think of the connective

‘if…then’, as being a referential representation which represents the conditional as a relation between two representations. We can then think of the referents of this relation as ordered pairs of representations, which are referents of ⌜if…then⌝ just in case the first representation of the set is false, or both representations are true. If this is the right reading of the connective ‘if…then’ in the sentence above, we have to take the two sentences “Peter is grey-haired”, and “Peter is old” to themselves express referential representations which refer to truth-evaluable representations; those truth-evaluable representations that would be expressed by the sentences if used by themselves.

This reading, however, has the drawback that while we may use a sentence as a name for the representation it would express if used by itself, it is not what we commonly seem to do with a sentence. We do not usually use sentences as referential representations. So the standard reading of the above sentence - that the logical vocabulary is used to indicate the relationship which needs to hold between the referents of the constituent referential representations for the representation to be true - is preferable to the reading on which the sentences act as names for the representations expressed.

There is an interesting connection between the two proposals for reading the sentence, however. The representation consisting of the relation ⌜if…then⌝ and the names for the truth-evaluable representations, and the complex representation consisting of ⌜Peter⌝,⌜old⌝, ⌜Peter⌝ and ⌜grey-haired⌝ are bound to have the same truth value. This is because the reference realizing properties for the referential representations referring to the truth-evaluable representations (⌜Peter is grey-haired⌝ and ⌜Peter is

old⌝) are, we may suppose, that the referents have the requirements for truth the representation would have if the sentence was used to express a truth-evaluable representation, and that the reference realizing properties of the constituents are the same as the ones the constituents would have, if the sentence was used as a truth-evaluable representation. So the representation ⌜if…then⌝ would refer to the ordered set of representations named by the two sentences, just in case the requirements for truth of the representations expressed on the first reading are satisfied.

All this, of course, only applies under the assumption that we use the sentences as names for truth-evaluable representations such that they refer to the truth-evaluable representation they would express, if used independently.

Be that as it may, however, I take the proper reading of the example sentence to be that it expresses a truth-evaluable representation consisting of the referential representations ⌜Peter⌝,⌜old⌝, ⌜Peter⌝ and ⌜grey-haired⌝, and as having the suggested complex requirements for truth outlined above.

The discussion illustrates how a complex constitution of truth-evaluable representations, for example through the use of logical vocabulary, gives rise to complex requirements for truth. So the complexity of representations can be captured with the concepts developed and is not an obstacle in the use of the present framework in an explanation of necessary truth.

6. The Case for Representations

The semantic framework with representations as the primary objects of study has some advantages over the classical study of

language as based on the communication model when investigating truth, reference, and the modes of truth.

The first and salient advantage is that the focus on representations generally allows for a general account of necessary truth. For truth is a property that applies not only to sentences, if used as representations, but also to other kinds of representations, for example mental states. Consequently, the modes of truth, necessity and possibility, also apply to all these different kinds of representations. So if there is any explanation to be had for the modes of truth which proceeds from properties of the representations, then these should better not be the properties only one kind of representation, for example sentences, have. Rather, the focus has to be on the representational properties the different kinds of representations share. Only by considering properties all kinds of representations have, can there be a general explanation of necessary truth in terms of the features of the representation. This is the first reason to focus not on sentences of a language independently from their use as representations, but to look at the objects which have their representational properties essentially.

Second, even if we can identify some mechanisms by which a sentence cannot fail to express something true, this will not suffice for the necessary truth of the representation expressed. For a necessary truth is not necessary, because the sentence expressing it cannot but express a true proposition, but because what it says, its content, cannot fail to be the case. Thus only objects which represent what they do essentially, that is,

representations in the sense introduced above, can be properly said to be true necessarily.[112]

Words and sentences are contingently associated with concepts and propositions.[113] Thus the association between words and concepts, and sentences and propositions can and does change, consequently what these words or sentences mean can change.[114] The word 'bachelor', for example, may presently be associated with the complex concept 'unmarried and male', but this may change, for example, if it starts to include more people on the gender spectrum, and not only men. Thus, while it may today still be true to say "all bachelors are men", in the future the same words may be used to state a falsehood. Thus, sentences, even if they are not context-sensitive, only have fixed truth-conditions at or during a time, and can thus not be said to be true or false simpliciter. To deal with this problem we could talk about truth at a time, and necessary truth at a time, but the time index would not be explanatory of why the truth-values of the sentences change. What accounts for the change in truth value of a sentence is a

---

[112] Another candidate for being the bearers of necessary truth are propositions in the traditional sense. However, traditional propositions preclude an explanation of their necessary truth, because they lack the relevant features that can be exploited for an explanation of the necessary truth of truth-evaluable representations.

[113] As above, I leave open what propositions and concepts are. On an intuitive communication model, they are just mental representations, but on virtually all accounts prominent today, they are abstract entities of some kind: Sets of worlds, sets of properties and individuals, etc.

[114] It is quite controversial how words and sentences, as the objects of study in linguistics, should be individuated, and how independent their individuation is of their meaning. For discussion see Kaplan (1990), Hawthorne and Lepore (2011), and most recently Bromberger (2011). What is uncontroversial, however, is that words and sentences are only contingently associated with their full meaning, and that their meaning may change over time.

change in the representational properties of that sentence at different points in time.

So talking about the objects which have the relevant representational properties essentially instead of those that have them only accidentally, is the best and surest way to secure that the object is the proper bearer of truth and necessary truth simpliciter.

I conclude that representations are the proper and reasonable objects of study for an account of necessary truth which seeks to explain this necessary truth by the representational properties of the representation. Before finally moving on to the envisaged explanation of necessary truth, I want to clarify the relationship between the present approach to representations and linguistic research.

The communication model, linguistic evidence, and representations

The discussion above emphasized the difference between the communication-model and the representation-framework. There are, however, connections between the two approaches, and at no point should the present suggestion be understood as an attack on the valuable work of linguists working within the communication paradigm. In this last section of the chapter, an important contribution linguistic evidence can make to the investigation of representations is highlighted, but first some important differences are pointed out.

Linguists focus on one representational system in particular: language as we use it in communication. The interests in this phenomenon are quite diverse, but one central branch of linguistics is concerned with analyzing how words and phrases of

a language and the way they are put together, contribute systematically to the truth-conditions of sentences.[115] The constituent words or phrases are assigned an interpretation (often just its extension, but depending on the level of analysis also its intension) and it is then shown how the words with their interpretation contribute to the truth values of sentences. Taking an example from an introductory textbook, the sentence "Ann smokes", can be analyzed into the noun-phrase 'Ann', and the verb-phrase 'smokes', which are exemplified by the noun 'Ann', and the verb 'smokes'. The noun is assigned a denotation: Ann. And the verb is assigned a denotation: the function which returns the truth value 'true' for each elements x of the domain, iff, x smokes. The denotation of the entire sentence is a truth-value, and since the structure is as it is, it is the truth value which the denotation of the verb-phrase returns, if the denotation of the noun-phrase is taken as input. That is, it returns the truth value 'true', iff Ann smokes.[116] This rather simple example gives a taste of how a certain branch of linguistic analysis proceeds, and it highlights how linguistic research is helpful for the present project, but also where the differences lie.

First, the differences. Linguistics is interested in the workings of language or languages as systems of communication that transmit information and are used by human beings. As such linguistics focuses on words and sentences, and how the same word or sentence type is regularly used to convey the same informational content. Since the meaning of words plays a systematic role in

---

[115] Truth-conditions are important to the linguist, because sameness of truth conditions is a good indicator for sameness of meaning, even if truth-conditions cannot be identified with meaning.

[116] The example is taken from Heim and Kratzer (1998) p. 13ff.

constituting what the informational content of a sentence is, the contribution of the meanings of words to the meanings of sentences are also studied. Thus linguists are interested in words of a language, and how they are used in communication to represent something, not so much in representations themselves.

The most important difference between the study of language in communication, and the study of representations lies in the role reference and truth play in the investigation. Consider again the toy example above. Here the verb-phrase is construed as a function from objects to truth-values, and its extension, its reference, is conceived of as the objects for which the function yield the value 'true'. Thus reference is secondary to truth in the order of inquiry. This is useful for a linguistic analysis for two reasons: First, it makes intelligible how the verb-phrase contributes to the truth-value of the whole sentence, namely by taking as input the denotation of the noun-phrase to yield a truth-value. Second, investigating truth-value changes in response to variations in input is a good method for distinguishing different functions, and thus

for investigating the meaning of these words, if this meaning is conceived of as the contribution of a word to truth-conditions.[117]

Useful and adequate as this outlook is for linguistic research, conceiving of phrases in predicate-position as functions contributing to truth-conditions is putting the cart before the horse from the point of view of the representation framework, which

---

[117] As is well known, this method works quite well for most cases, but fails for (necessarily) coextensive predicates with different meanings. Thus, meanings cannot be identified either with extensions or with intensions, and consequently predicates are not purely functions from domains to truth values. For most purposes, however, truth-conditions are a very good guide to meaning, and so formal modeling often works within the truth-functional paradigm.

seeks to explain how reference and truth, constitutively, come about. For it is not in virtue of it being true of everybody who smokes that they are referents of the representation ⌜smokes⌝, rather it is in virtue of them being referents of ⌜smokes⌝ that it is true to represent them as smoking. In a nutshell, linguistic research takes truth and reference for granted, and uses these concepts to find out about the systematic use of words and sentences in communication. The representation-framework, on the other hand, is interested in different objects, representations, and in what constitutes their reference.

Even though the basic outlook is quite different, linguistics, as special science of one kind of representation, can be useful for the study of representations. For by investigating the semantic assignments, it contributes to finding out what a particular sentence is used to represent in some context. That is, it helps us find out which parts of a representation are referential, to what they refer, and how the other vocabulary indicates what the requirements for truth are. So once we consider sentences and words as concrete examples of representations, linguistic insights can help deliver part of the evidence needed for a correct account of the representational properties of particular representations.

Summing up

In the present chapter the framework and the basic concepts for the explanation of necessary truth were developed. I argued that if there is to be any chance for giving an explanation of necessary truth in terms of the representational properties of the constituents of representations, we cannot focus on words and sentences, for they are the wrong kind of object to be true or false simpliciter.

Rather we have to look to representations and the central representational relations reference and truth.

Representations have certain representational properties essentially. We can distinguish between referential representations and truth-evaluable representations. Referential representations can be divided into singular and general. The former refer to one and only one object, while the latter, at least potentially, refer to a number of objects. A reference relation between a referential representation and a referent is established in virtue of properties of the representation and properties of the referent. The former are the reference determining properties, while the latter are the reference realizing properties. Which properties of the referent the reference realizing properties are, is determined by the reference determining properties of the representation. The reference realizing properties play a special role in the individuation of referential representations: Token representations are of a kind, if and only if they determine the same reference realizing properties.

Truth-evaluable representations are constituted by referential representations and their truth depends on the relationship between the referents of the constituent referential representations. The relationship required for the truth of a truth-evaluable representation are the requirements for truth of the representation. Which relationship is required for the truth depends on how the truth-evaluable representation is constituted by the referential representations. Truth-evaluable representations can be individuated with respect to their constituents and their requirements for truth: Truth-evaluable representations are of a kind, if and only if they have, as constituents, referential representations of the same type, and have the same requirements for truth.

Within this framework, I will argue, necessary truth can be explained. It is the task of the next chapter to show how this may be done.

Chapter 5: An Explanation of Necessity

The Kant-inspired explanation of analyticity is an explanation of how some judgements cannot fail to be true in virtue of their constituents standing in the appropriate containment relations.[118] The main fault with this explanation, if we take it to explain necessity generally, is that it is not nearly general enough to explain the interesting cases of necessary truths: Only few concepts seem to exhibit a containment structure and so an explanation of necessity that relies on this structure cannot work for the many judgments whose constituent concepts do not have it. But the way in which Kant explains how analytic judgments cannot fail to be true, can serve as a blueprint for an explanation of necessity, because giving an explanation for the fact that a judgement cannot fail to be true amounts to giving an explanation of the judgment's necessity. The general account of representations and their semantic relations developed in the last section can be utilized to give a general explanation of some representation's necessary truth by explaining how they cannot fail to be true in virtue of how the reference of their constituents is determined. Situating the explanation in the more general framework will show that the containment structure for concepts is merely a superficial feature of the explanation of why analytic

[118] It should be noted again that it was not Kant's goal to explain necessity. His main concern was to explain how synthetic *a priori* knowledge is possible, while analytic judgments, supposedly *a priori* and necessary, are just the boring cousin of these more interesting truths. Consequently, the criticism should not be read as directed at him. Kant knew about this feature of his definition of analyticity and endorsed it. I am not concerned with criticizing or supporting the Kantian project, I merely take Kant's definition as a starting point for what I take to be a good explanation for necessary truth.

judgments cannot be false, and that the necessity of a number of philosophically interesting truths can be explained in this way.
First, the Kant-inspired explanation will be rehearsed once more to clarify what is doing the explanatory work. Second, the explanatory strategy will be used to show how the necessary truth of representations can be explained more generally with the help of the semantic framework developed. In the next chapter, I will argue that this explanation of necessity covers a range of philosophically interesting examples of necessary truths.

1. Kant's Definition of Analyticity and the Explanation of Necessity

Consider again the Kantian definition of analyticity:

> "In all judgments in which the relation of a subject to the predicate is thought […] this relation is possible in two different ways. Either the predicate *B* belongs to the subject *A* as something that is (covertly) contained in this concept *A*; or *B* lies entirely outside the concept *A*, though to be sure it stands in connection with it. In the first case, I call the judgment analytic, in the second synthetic."[119]

The definition of analyticity itself does not tell us much about why concept-containment gives rise to necessity, it merely tells us that judgments with subject-predicate structure are analytic: if the predicate-concept is contained in the subject concept. Explanation, however, is not far to seek. The first thing to note is that concepts should quite literally be thought of as having other

[119] Kant KrV A 6-7.Translation from Kant (1998).

concepts as parts: Concepts build up other concepts.[120] The reference a built-up concept has is determined by its constituent concepts. So for example, if a concept is built up from the two concepts 'red' and 'ball', then its referents are all things that are the referents of both concepts 'red' and 'ball'.[121] The second thing to note is the required subject-predicate structure of the analytic judgment. This structure leads to the following relationship between reference and truth: For a judgment with subject-predicate structure to be true, the referents of the subject concept must also be referents of the predicate concept.

The containment of the predicate concept in the subject concept as a part determines the reference of the subject concept to be such that its referents are also referents of the predicate concept. This in turn guarantees that the judgment is true, since the truth of a subject-predicate judgement requires the referents of the subject to also be referents of the predicate. So the concept structure, the judgment structure, and the resulting constraints on reference and truth work together to explain why an analytic judgment cannot fail to be true: It cannot fail to be true, because the reference of the subject and the predicate concept is determined in such a way that its requirements for truth are guaranteed to be satisfied.

While Kant's definition of analyticity heavily relies on containment and the subject-predicate structure of analytic judgments, the explanation of why these judgments cannot fail to be true does not essentially rely on these two elements. What explains that such a judgment cannot fail to be true, is the

---

[120] Hence the name 'analytic'. For a historically informed discussion of parthood and concepts in Kant, see Leech (forthcoming 2016)

[121] In principle, there may be other ways in which concepts are built up which may give rise to different conditions on the referents, but apparently the Kantian containment-picture only takes account of the kind of constitution outlined here.

working together of how the reference of the constituents of the judgment is determined and what is required of the referents for the truth of the judgment. Containment of one concept in another, given the background theory of concepts, is a sufficient condition for the necessary truth of a subject predicate judgment. However, it is likely to not be a necessary condition, as there are possibly other ways in which reference determination of constituent representations guarantees the truth of truth-evaluable representations. We may call the Kantian analytic truths narrowly analytic to mark the distinction between those truths classically conceived of as analytic, and the wide analytic truths, which are necessary in virtue of how reference is determined generally. In what follows, the explanation of necessary truth will be expanded to cover not only the narrow analytic truths, but also the wide analytic truths.

2. Explaining Necessary truth

An explanation of why a truth cannot fail to be true amounts to an explanation of its necessity. The explanation suggested by the Kantian account of analyticity goes like this: An analytic truth cannot fail to be true, because the relationship of the referents of its constituents required for its truth is guaranteed by how the reference of the constituents is determined. This explanation cannot only be applied to narrowly analytic truths whose constituents stand in containment relations to each other, but to other representations as well, provided that something about how the reference of the constituents is determined guarantees that the relationship between the referents required for the truth of the representation obtains.

The account of representations outlined in the last section provides all the tools needed for a general account of necessary truth. The starting point for an account of necessary truths are token representations. Token truth-evaluable representations are constituted by token referential representations. The truth of truth-evaluable representations depends on the relationship between the referents of the referential representation. The relationship the referents need to have for a truth-evaluable representation to be true are the referential requirements for truth, which can be stated in the following form:

> *Requirements for truth:* The truth-evaluable representation R is true, iff the referents of the constituent referential representations C1,…,Cn stand in relation I.

The constituent token representations refer to something in virtue of their reference determining properties, and in virtue of the reference realizing properties of the referents. The reference determining properties determine what the reference realizing properties, the properties the referent(s) of the token representation must have for a reference relation to hold between the representation and its referent, are. The reference realizing properties can be stated as conditions. They take the following form:

> *Reference conditions:* The referential representation C refers to o, iff o has P1,…,Pn.[122]

[122] P1,…,Pn are the reference realizing properties.

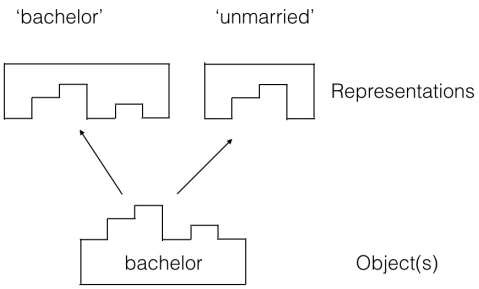
These conditions are the reference conditions of representation C. The reference conditions typically state necessary and sufficient conditions for reference to occur. However, I will also use the term reference conditions to refer to the merely necessary conditions.[123]

For a truth-evaluable representation to be necessarily true, the reference realizing properties of the constituent representation must be appropriately related such as to guarantee the satisfaction of the referential requirements for truth. One way in which this may happen is if the requirements for truth require the referents of one constituent representation (C1) to be a subset of the referents of the other (C2), and if the reference realizing properties of C2 are a subset of the reference realizing properties of C1. One example of such a case is the example of the unmarried bachelors above. ⌜All bachelors are unmarried⌝ cannot fail to be true, because the requirements for truth require the referents of ⌜bachelor⌝ to be a subset of the referents of ⌜unmarried⌝, and the reference realizing properties of ⌜unmarried⌝ are a subset of the reference realizing properties of ⌜bachelor⌝, since we can assume, along with standard treatment of the conditions the referents of ⌜bachelor⌝ have to satisfy, that these referents are unmarried.

---

[123] These reference conditions may be disjunctive. There is no requirement that the conditions need to be stateable as singly necessary and jointly sufficient conditions, although this is desirable. There may be limits to the properties expressible in a language, so it should not be expected that it is always possible to spell out the reference conditions for every representation as sufficiently specific necessary and sufficient conditions.

The relationship between the reference determining, and the reference realizing properties can again be illustrated with the model of fit introduced in the last section.



The picture illustrates that the reference determining properties determine the reference realizing properties of ⌜bachelor⌝ and ⌜unmarried⌝ such that the reference realizing properties of ⌜unmarried⌝ are a subset of the reference realizing properties of ⌜bachelor⌝.

This is the simplest case of how the reference realizing properties can be related to the requirements for truth. More complex representations have more complex requirements for truth and accordingly more complex relations between the reference realizing properties that need to be in place for their satisfaction to be guaranteed.

To capture all these cases, I propose the following definition:

> *Necessary truth (tokens):* A token truth-evaluable representation is *necessarily true*, iff the satisfaction of its requirements for truth is implied by the reference conditions of its constituent referential representations.

The definition can be applied to our example in the following way: The requirements for truth of ⌜all bachelors are unmarried⌝ are: ⌜all bachelors are unmarried⌝ is true, iff the referents of ⌜bachelor⌝ are referents of ⌜unmarried⌝.

The reference conditions of ⌜bachelor⌝ are, we may assume: ⌜bachelor⌝ refers to o, iff o is unmarried and male.

And the reference conditions for ⌜unmarried⌝ are, we may assume for simplicity: ⌜unmarried⌝ refers to o, iff o is unmarried.[124]

It follows from these reference conditions that the referents of ⌜bachelor⌝ are referents of ⌜unmarried⌝ and consequently the requirements for truth of ⌜all bachelors are unmarried⌝ are satisfied.

The definition of necessary truths in terms of reference conditions and what they imply is adequate in light of what reference conditions and requirements for truth are. However, the present definition is neither intended to be a reduction, for it appeals to the modal notion of entailment, nor does the definition give the

---

[124] These reference conditions are trivial, but this triviality is not harmful to the example, as was explained in the last chapter.

desired explanation of necessary truth, even though it points to an explanation.

The intended explanation of necessary truth involves the reference determining properties, the corresponding reference realizing properties, as well as the requirements for truth of the truth-evaluable representation. Let us look at each ingredient to the explanation in turn. The fact that a representation's reference determining properties determine reference realizing properties is gained from considering the nature of reference. Since reference is a relation between a representation and an object or a number of objects, which holds in virtue of properties of the representation and properties of the referent, the referent(s) are bound to have some reference realizing properties determined by the reference determining properties. The fact that truth-evaluable representations are true in virtue of the referents of its constituents standing in some specified relation to each other is equally an uncontroversial fact about truth-evaluable representations. Now it may be the case that the relationship between the referents of the constituent representations that needs to hold for the truth-evaluable representation to be true, is guaranteed to hold in virtue of how the reference realizing properties relate to each other. If this is the case, the truth-evaluable representation cannot be false, and so it is necessarily true. In a sentence:

> *Explanation of Necessity:* A necessarily true truth-evaluable representation is necessary, because the reference determining properties of the constituent referential representations determine the reference realizing properties such that the requirements for truth of the truth-evaluable representation are guaranteed to be satisfied.

What it takes for a representation to be necessarily true, is that its requirements for truth are guaranteed to be satisfied. That the conditions for this guarantee are in place can be gained from the relationship between the reference realizing properties of the constituent representations and the requirements for truth of the truth-evaluable representation.

This explanation may at first glance seem circular, because it involves a "guarantee" that the requirements are satisfied, which may seem like nothing more than a restatement of the necessary truth of the representation. But the present account is not intended as a conceptual reduction, but as an explanation, and it explains how the guarantee is effected. For each (kind of) case, we can say without recourse to the modal notion of guarantee, what it takes for the requirements for truth to be guaranteed by the reference realizing properties.

Consider again the example of the unmarried bachelors. The representation ⌜all bachelors are unmarried⌝ is true, iff the referents of ⌜bachelor⌝ are a subset of the referents of ⌜unmarried⌝. And thus it is necessary, according to the present account, just in case the reference realizing properties determined by ⌜bachelor⌝ include among them the reference realizing properties determined by ⌜unmarried⌝.[125] This way of saying what it takes for the particular representation to be necessarily true, makes no reference at all to any modal notions, and thus the characterization is not explanatorily circular.

---

[125] Note that the inclusion relation between these sets of properties is not the same as the relation of concept containment in the Kantian account. For concept containment, if it exists, is only one possible way that gives rise to this relationship between the reference realizing properties.

The apparent circularity comes in when we try to say more generally what it takes for a representation to be necessary, for there are different ways, depending on the requirements for truth, in which the relationship between the reference realizing properties can guarantee the satisfaction of the requirements for truth.[126] So one has to revert to a description in modal terms: what it is to be necessary is to be guaranteed to be true. This, however, is unobjectionable, for the intention is not a conceptual reduction, a different way of saying the same thing, but constitutive explanation. So as long as we can explain what gives rise to such a guarantee in each particular case without circularity, the explanation is not objectionably circular. We can do this in each case by giving the relation between the reference realizing properties with respect to the relation between the referents required for the truth of the representation which explains the necessary truth.

Here are three more examples of how the guarantee may be effected in some cases. Lets assume for simplicity the following format for the examples: the truth-evaluable representation R consist of two constituent referential representations C1 and C2, which determine certain reference realizing properties. The requirements for truth for R are given as a relation between the referents of C1 and C2.

---

[126] One may attempt a list of different kinds of truth-evaluable representations, which classifies them according to the relationship the referents of their constituents need to stand in. Then one could, for each such kind, say more generally, without recourse to modality, what the required relationship between the reference realizing properties needs to be. There is no need for this here, however, as the modal vocabulary will do just fine to get the intended explanation across.

The first example is a representation which requires for its truth that the referents of C1 are a subset of the referents of C2. A representation with this form of requirements for truth is necessarily true, just in case the reference realizing properties determined by C2 are a subset of the reference realizing properties determined by C1.

The second example is a representations which requires for its truth that the referents of C1 are identical to the referents of C2. A representation with this form of requirements for truth is necessarily true, just in case the reference realizing properties determined by C1 are the same as the reference realizing properties determined by C2. That is, it requires that both representations are of the same type.

The third example is a representation which requires for its truth that the referents of C1 and C2 are disjoint. A representation with this form of requirements for truth is necessarily true, just in case either the reference realizing properties of C1 include the negation of the reference realizing properties of C2, or vice versa.

What explains necessity in all three cases is the appropriate relationship between the requirements for truth and the reference realizing properties, a relationship that guarantees the satisfaction of the requirements for truth.

The proposed explanation of necessary truth shows how representations can be guaranteed to be true in virtue of having an appropriate relationship between the relevant reference realizing properties and the requirements for truth. But one may wonder what the scope of the account is, whether it captures all philosophically interesting cases of necessary truths, or whether it is just a more abstract way of accounting for mere conceptual truths. This is the challenge that will be confronted in the next

chapters by giving a range of examples and by arguing that the kind of necessity explained is indeed metaphysical necessity.

The purpose of this section was to show that an explanation of why some representations cannot fail to be true naturally drops out of a general account of representations, given that reference is not primitive, and given that the truth of truth-evaluable representations depends on the reference of its constituents. I take these two assumptions to be uncontroversial and thus that it has been established that at least some kind of necessary truth can be properly explained in the way advocated.

3. Necessary Truth for Types of Representations

The explanation of necessity was given in terms of token representations, since they are naturally taken to be the primary bearers of truth and necessary truth in the present framework. The explanation, however, can be extended to types of token representations. For simplicity I will sometimes call types of truth-evaluable representations 'propositions'. And I will sometimes call types of referential representations 'concepts'. However, it should be kept in mind that this is a non-standard and stipulative use of both terms, and should not be taken to be the standard usage of the term proposition or concept as it is used in most philosophical debates.[127]

[127] There is not *the* standard use of these terms, but the debate around propositions fits better with the communication model discussed above, where propositions and concepts have a very specific theoretical role, that the concepts and propositions as they are used here need not satisfy. Still, conceiving of propositions as some type of a token representations is suggested by some philosophers. Scott Soames in King, Soames, Speaks, (2014), for example, argues that one should conceive of propositions as a type of act of predication.

Truth-evaluable representation tokens are of a type in virtue of having constituent representations of the same type, and in virtue of having the same requirements for truth, so we may say that propositions have derivative requirements for truth. Types of referential representations (concepts), accordingly have derivative reference realizing properties and thus derivative reference conditions, for the token referential representations are of a type in virtue of having the same reference realizing properties. The definition for the necessity of propositions is thus parallel to the definition of necessary truth for token representations:

> *Necessary truth (types):* A proposition is *necessarily true*, iff the satisfaction of its (derivative) referential requirements for truth is implied by the (derivative) reference conditions of its constituent concepts.

As before, the definition is not a reduction and should not be confused with the explanation of necessity. The explanation of necessary truth runs via the interplay between the reference determining properties of the token constituent representations and the token truth-evaluable representation's requirements for truth.
We can take the definition of necessity for propositions to tell us what the reference determining properties of the tokens of the constituent concepts and the requirements for truth of the token truth-evaluable representation would have to be like for them to necessarily represent truly.

4. Adequacy of the explanation

The explanation of necessity, I want to argue here, has good prospects for satisfying the criteria of adequacy for a philosophical explanation outlined in part one: non-circularity, generality, and intelligibility. I hope to convince you throughout the thesis that it does so better than its competitors. In the following, I consider the criteria of adequacy in turn, and indicate where further argument is needed for an adequate assessment of the merits of the theory.

The non-circularity of the explanation was discussed above already. Although some apparently modal notions appear in both, the definition, as well as the general formulation of the explanation, this doesn't threaten the non-circularity of the explanation. To see why the explanation is non-circular, it helps to look back at the discussion in the first part on what a philosophical explanation is supposed to do: It is supposed to tell us what it takes, constitutively, for something to be the case. An explanation of knowledge, for example, is supposed to tell us what it takes to be knowledge, and an explanation of necessary truth is supposed to tell us what it takes to be necessarily true. This question can be approached in different ways. One way is to give a conceptual analysis of a philosophical concept, that is, essentially, to find a different way of representing the intended philosophical concept, and to hope that this reformulation elucidates what it takes for the philosophical phenomenon to obtain. As is well known, however, such analyses are not always possible, and even if some extensionally adequate reformulation is possible, it is not always clear that it moves us any closer to knowing what it takes for the phenomenon to obtain. In particular in the modal case, when we want an explanation of why some

representation is necessarily true, it seems that we cannot expect a (sensible) reformulation that is not itself in need of further explanation.[128] Consequently, the present theory did not attempt any such conceptual analysis. Rather, a reformulation was chosen that is in further need of explanation, but which moves us closer to understanding what it takes for a representation to be necessarily true: necessary truth is a guarantee of truth. Subsequently, the need for explanation was discharged by showing what it takes for a representation to be guaranteed to be true. Something must make it the case that the representation cannot fail to represent truly. The way the guarantee is effected, I argued, is via the appropriate relationship between what is required for the truth of the truth-evaluable representation and the reference realizing properties determined by its constituent representations. What this relationship is, is different for different kinds of truth-evaluable representations, and so there is no way of saying what the relationship for any truth-evaluable representation is without resorting to the modal notion of a guarantee. However, even though the notion of guarantee appears in the official explanation of necessity, the explanation is *not* that a representation is necessarily true, because it is guaranteed to be true. This would rightly count as a circular explanation. Rather, the explanation is that what it is (not takes) for a representation to be necessarily true, is to be guaranteed to be true, and what it takes for this guarantee to be effected is an appropriate

[128] Lewis' genuine modal realism would be an account that succeeds in a reformulation that does not seem to be in immediate need for further philosophical explanation (Lewis 1986). However, to even be extensionally adequate, it requires the existence of a plurality of concrete possible worlds, and even if all these worlds existed, it seems like quite a stretch to claim that truth in all these worlds is really what it takes for a representation to be necessarily true.

relationship between the requirements for truth and the reference realizing properties. What the appropriate relationship for some specific kinds of truth-evaluable representations is, was outlined in the examples. So while I agree that there is no way to give an illuminating conceptual reduction of necessity, I do hold that an illuminating and non-circular explanation of why each necessary truth is necessary can be given.

Whether the second requirement, generality, is satisfied remains to be shown in the following sections, but presently it is useful to see what would have to be the case for the requirement to be satisfied. The generality requirement states that a philosophical explanation should be able to not only give a partial account of some phenomenon, but a general and unified account of an apparently unified phenomenon: necessary truth. To argue that the requirement is satisfied, I will first have to show that the account is extensionally adequate, that is, that it is able to explain uncontroversial examples of metaphysically necessary truths.

I will argue for the extensional adequacy of the theory in two ways. In the next chapter I will give examples of standard necessary truths and argue that they can be explained in the way envisaged. In the following chapters I will provide additional considerations that are designed to strengthen the claim to extensional adequacy. The examples will suggest that the necessity primarily explained by the account is what has commonly been called metaphysical necessity. However, since some possibly controversial assumptions about the reference realizing properties must be made for the examples to go through, there is room for some doubts about extensional adequacy. Some of these doubts will be alleviated by comparing the present theory to essentialism, a theory which is commonly taken to be

extensionally adequate, if it explains necessity at all. I will suggest that we should identify essential properties with those reference realizing properties of an object which are determined by the singular representation referring to it. If this identification is plausible, then it speaks strongly in favor of the theory giving a general explanation of metaphysical necessity.

Since metaphysical necessity is not the only kind of necessity the theory may be expected to explain, but also other, weaker kinds of (alethic) necessity, such as nomic or practical necessity, chapter seven is devoted to arguing that the account can be appropriately expanded to account for these kinds of necessity as well. Given that these considerations turn out to be convincing, the proposed explanation has a strong claim to being a sufficiently general theory of necessary truth.

I also take the proposed explanation to score high on the last desideratum: intelligibility of the explanans. For the explanans makes use of uncontroversial resources and does not require us to believe in anything we would usually reject. The resources merely rely on very general considerations about the nature of representation and reference: (i) that truth-evaluable representations have as their constituents referential representations, (ii) that they are true, if the relationship between the referents of these constituent representations is as required, and (iii) that reference is a relation that holds in virtue of properties of the referent and the representation. Unless one is a skeptic about reference and truth, these assumptions should not be controversial, for no matter what philosophical view one is inclined to have on how reference is determined, it is quite uncontroversial that for the reference relation to hold, both the representation as well as the referent must have certain properties.

Equally, if one accepts that there is reference, the proposed relationship between reference and truth can hardly be doubted. So necessary truth is bound to arise from quite basic assumptions about how representations work.

These considerations provide reason for optimism that the present explanation is a non-circular, general, and genuinely intelligible theory of necessity. It will have to be shown in the following chapters that this optimism is justified.

Chapter 6: Applying the Theory

The last chapter outlined how an explanation for the necessary truth of representations may go. In this chapter the explanatory strategy will be applied to examples. This application is supposed to clarify the explanation itself, but it is also supposed to show that it explains what has commonly been called metaphysical necessity. To do so, I will apply the theory to standard examples of metaphysically necessary truths, and argue that their necessity is indeed explainable in the way suggested.

Using examples, however, can only go part of the way of showing that the theory is (extensionally) adequate. This is for two reasons. The first is the not terribly interesting reason that examples are just that, examples, and the small number given here can only raise our confidence in the theory so much. However, since there is reasonable hope that the account covers more examples, I will for now pay no further attention to this problem. The second reason for doubting the effectiveness of the examples for showing that the theory has a claim to extensional adequacy, however, is more worrying: I will have to make assumptions about what the sentences, commonly taken to be used to express necessary truths, are used to represent. That is, I will have to make assumptions about the requirements for truth of the representations, and I will have to make assumptions about the reference realizing properties of the constituent representations. These assumptions will not just be made up, but they will also not be beyond doubt. I will take care to give plausible reference realizing properties and requirements for truth in light of my understanding of the english sentences said to express a necessary truth, and in light of what we know about the reference realizing

properties of the constituent referential representations. While this will provide some reason to endorse the explanation as extensionally adequate, this is conditional on the plausibility of the assumptions. Thus I will not merely rely on the examples to argue for the extensional adequacy of the theory, but in the following chapters give some more reasons to take the present explanation to explain metaphysical necessity.

In the application of the theory, I will proceed as follows. First, I give examples of representations whose necessity is due to the reference realizing properties of general representation.[129] These examples are probably the most uncontroversial examples, as some of them are even today taken to be conceptual, or analytic truths in some sense. The examples that follow will show that the difference between these apparently simpler examples to necessary truths, whose necessity can be traced to the reference realizing properties of the singular representation, does not run deep. Necessarily true representations, where the necessity can be traced to the singular representations, do have a certain peculiarity, however. This is due to the apparently existence-entailing character of sentences involving a singular representation in subject-position. For if such sentences are

---

[129] I talk here of necessary truths whose necessity is due either to a general or a singular representation. Of course the full explanation of necessary truth involves the reference realizing properties of all referential representations in the truth-evaluable representation and their relation to the requirements for truth. Still there sometimes is a sense in which the necessity is due to one of the referential representations. This is vivid for cases where the reference realizing properties of one representation are a subset of the other. Here the representation with the larger set of reference realizing properties can be conceived of as the one 'responsible' for the necessary truth. As an example take once more ⌜all bachelors are unmarried⌝, where the necessity may be said to be due to the representation ⌜bachelor⌝.

indeed used as truth-evaluable representations which require for their truth the existence of a referent, then it seems that the present theory cannot account for their necessity, because there is no way to guarantee that anything in fact has the reference realizing properties determined by the reference determining properties of the representation. Luckily, however, there is good reason not to take these sentences to be used as representations which require the existence of the referent of the singular representation for their truth.

To argue for all this, I start with necessarily true representations involving general representations. This is followed by an interlude discussing the problem of existence, necessary truth, and singular representations generally. The insights gained there are then be applied to examples of necessarily true representations involving singular representations.

1. Necessarily True Representations Involving General Representations

Conceptual truths

The first example is one we have already encountered above and an example of what is sometimes called a conceptual truth.[130] The example is the representation expressed by the sentence "Vixens are female".

The first step in explaining the representation's necessity, is to identify the referential representations and the requirements for

[130] From the discussion of analyticity in the last part it should have become clear that I do not take conceptual or analytic truths to be made true in any special way. The phrase "conceptual truth" is here merely used to classify the class of truths commonly called "conceptual truths" by philosophers. Similar remarks apply to the labels "logical truth" and "mathematical truth".

truth of the truth-evaluable representation: The representation ⌜vixens are female⌝ should be taken to consist of the two referential representations ⌜vixen⌝ and ⌜female⌝. The subject predicate structure of the sentence also gives us a clue as to which relationship is required for the truth of the truth-evaluable representation: The referents of ⌜vixen⌝ must be referents of ⌜female⌝.

The next step in the explanation of necessary truth is identifying the appropriate relationship between the reference realizing properties determined by the referential representation which explains that the requirements for truth are guaranteed to be satisfied. As I pointed out in the last chapter, the relationship between the reference realizing properties explaining the necessary truth is the following: The reference realizing properties determined by ⌜female⌝ must be among the reference realizing properties determined by ⌜vixen⌝, for if this is the case, the representation ⌜female⌝ will refer to anything the representation ⌜vixen⌝ refers to, thus making sure that the requirements for truth are satisfied.

The final step in establishing the necessary truth of the representation is to find out whether the reference realizing properties determined by the representations are related in the way required for the representation to be necessarily true. To establish this, I will assume that the simple description theory of reference outlined in the semantic preliminaries is true for the representation ⌜vixen⌝. On this theory, the reference of ⌜vixen⌝ is determined by the appropriate association with other (token)

representations in the speaker's head. In the present case these are the token representations ⌜female⌝ and ⌜fox⌝. By being so associated, the reference realizing properties of ⌜vixen⌝ are determined to be the conjunction of the reference realizing properties of ⌜female⌝ and ⌜fox⌝. This is already almost enough to establish the necessary truth of the representation ⌜vixens are female⌝, all that has to be assumed in addition is that the referential representation ⌜female⌝ as it appears in the token truth-evaluable representation above is of the same type as the referential representation ⌜female⌝, which is associated with ⌜vixen⌝. If this assumption is true, then, since two token referential representations are of the same type in virtue of determining the same reference realizing properties, it is the case that the reference realizing properties of ⌜female⌝ are among the reference realizing properties of ⌜vixen⌝. And since this is just what is required for the representation to be necessarily true, it is necessarily true.

Note that the example did not even have to establish what the reference realizing properties are. It was sufficient to point out that they are appropriately related vis à vis the requirements for truth. That this is the case could be shown assuming a description theory of reference. It may be the case that the description theory presupposed here is not ultimately the best theory for explaining the reference of the representations. However, in the case of simple conceptual truths, it is at least plausible to assume the that reference of ⌜vixen⌝ suitably depends on the reference of ⌜female⌝.

The necessity of conceptual truths is thus explained by the theory. This, however, was to be expected since the blueprint for the explanation was the Kantian containment-account of analyticity, which could already explain these simple cases. So next it needs to be shown that the account covers more than these simple examples.

Logical truths

Logical truths are some of the prime examples of necessary truths and they should be accounted for by any theory of necessity. Standard examples are usually given in propositional logic: “p or not p”, “If p then p”, etc., I will first focus on these simple examples from propositional logic. Afterwards I will discuss examples from (first-order) quantified logic, as these need a slightly different treatment, due to the more complex character of the truth-evaluable representations.

Take first the simple example of “p or not p”. As it stands, it can hardly be taken to be a representation, more as representing the form of a representation, but by adding some detail and filling in assumptions, there is a reading of “p or not p” that may count as genuine representation. The propositional constants (p, q, r, …) can be taken to be referential representations that refer to propositions understood as truth-evaluable representations. The same letter is used as a representation of the same type on every occasion of use and it is used to refer to truth-evaluable representations of the same type. So each appearance of the letter ‘p’, for example, is to be taken to be a representation of the same type and to refer to truth-evaluable representations of type p. Multiple occurrences of propositional constants can be taken to

refer to ordered sets of propositions, depending on the order of their appearance.

The logical vocabulary (if…then, or, and, not) can then be taken to represent a relation between the representations which holds in virtue of their truth value. The referents of this relational representation are ordered sets of objects, in the present case, ordered sets of truth-evaluable representations.

The first step in establishing the necessary truth of the representation expressed by "p or not p", is to find out what its requirements for truth are. Given what has been said above, they are the following:

⌜p or not p⌝ is true, iff the referent of ⌜p…p⌝ is a referent of ⌜or not⌝.

The next step is to see whether the reference realizing properties of the representation are such that they guarantee the satisfaction of these requirements. Let us start with the general representation ⌜or not⌝. To be the referent of this representation an ordered pair of propositions must be such that either the first member of the pair is true and the second is true, or the first member of the pair is true and the second member false, or both members of the pair are false. The interesting part of these reference realizing properties is that it is sufficient for being the referent of the relation ⌜or not⌝ that both the first and the second element of the pair are either both true or both false. This will, for example, be the case whenever the requirements for truth of both representations of the pair are the same.

Next we need the reference realizing properties of the representation ⌜p…p⌝. Since this representation refers to an ordered set of truth-evaluable representations, the reference

realizing properties are that the referent is an ordered pair consisting of the referents of the representations ⌜p⌝ and ⌜p⌝. On the assumption that both ⌜p⌝ and ⌜p⌝, are representations of the same kind which refer to truth-evaluable representations of the same kind, the representation ⌜p…p⌝ refers to ordered pairs of truth-evaluable representations consisting of representations of the same type. Since representations are of the same type in virtue of having the same requirements for truth, it is a necessary condition for being the referent of ⌜p…p⌝ that the elements of the pair have the same requirements for truth.

This last finding is what we need to establish that the requirements for truth are guaranteed to be satisfied, for that both elements of the pair <p,p> have the same requirements for truth is sufficient for being the referent of ⌜or not⌝, and so the requirements for truth, that the referent of ⌜p…p⌝ is a referent of ⌜or not⌝ are bound to be satisfied.

The example shows how the necessary truth of logical truths of propositional logic can in principle be explained in the present framework, if the propositional constants are conceived of as representations of propositions. As it turns out, the explanation for their necessary truth is not very different from the explanation in the simple containment-cases.

Propositional logic contains only referential representations representing proposition, and relations that hold between them in virtue of their truth value. As such the referential representations only refer to other representations or sets of representations. Quantified logic expands this narrow focus, and adds more structure to the representations. By giving the representations

more structure, however, the function of the logical connective changes. They can no longer be viewed as representations of logical relations holding between propositions, as there are no longer representations of those propositions, rather, the logical connectives indicate the sometimes complex relationship between the referents of the referential representations required for the truth of the truth-evaluable representation.

Again an example involving quantification is helpful. Take as example the representation expressed by the sentence "For all x, if x is grey, then x is grey." The first step in finding out whether this representation is necessarily true is again to give the requirements for truth:

⌜For all x, if x is grey, then x is grey⌝ is true, iff the referents of ⌜grey1⌝ are referents of ⌜grey2⌝.

To distinguish the two token representations exemplified by the word 'grey', I have labeled them ⌜grey1⌝ and ⌜grey2⌝ respectively. The requirements for truth should in light of the standard reading of 'if…then' not be very surprising.

The next step is, as before, to see what the reference realizing properties determined by ⌜grey1⌝ and ⌜grey2⌝ are and whether they guarantee the satisfaction of the requirements for truth. We can assume that the representations ⌜grey1⌝ and ⌜grey2⌝ are of the same type, that is they determine the same reference realizing properties. In the context of some formal language this is obvious, as it is usually stipulated that if two predicates are symbolized by the same letter or word, then they are representations of the same type.

The argument that the reference conditions imply the satisfaction of the requirements for truth is now easily made: Since the

reference realizing properties determined by ⌜grey1⌝ and those determined by ⌜grey2⌝ are the same, a referent of ⌜grey1⌝ is a referent of ⌜grey2⌝. Since this just is the requirement for the truth of the truth-evaluable representation, the logical truth expressed by the sentence "For all x, if x is grey, then x is grey" is necessarily true.

Interestingly, logical truths do not look very different from other necessary truths in the present framework. What seems to be doing quite a bit of work to make some truths truths of logic is that some of the referential representations determine, by stipulation, the same reference realizing properties. This makes it easy to deduce the necessary truth of some truth-evaluable representations in which they occur, as no special attention has to be paid to what the exact reference realizing properties are.

Relations

Above a first special case of relations was discussed: logical relations holding between propositions. But other relations are ubiquitous and are also involved in necessary truths. So in the following I will give a non-logical example of a truth involving relations.

The special feature of relations is that they can hold between multiple objects and so the referents of relations cannot just be single objects, but can instead be conceived of as ordered tuples of objects. On this view, a two-place relation has as referent an ordered pair of objects, a three-place relation has an ordered triple of objects as referents, and so on.

The standard examples of necessary truths involving relations are, it turns out, quite complex, as they exploit the formal properties of

relations to form a necessary truth. Here is an example of a necessary truth involving relations; the representation expressed by the sentence “If Tobi is taller than Conny, and Conny is taller than Laila, then Tobi is taller than Laila.”

The structure of this truth-evaluable representation can be broken down in the following way: There are three general referential representations, each a representation of the dyadic relation taller-than. Second, there are three singular representations ⌜Tobi, Conny⌝, ⌜Conny, Laila⌝, and ⌜Tobi, Laila⌝ referring to the ordered pairs <Tobi,Conny>, <Conny, Laila>, and <Tobi, Laila>. Third, there are logical connectives, which indicate the requirements for truth.

The first step, as before, is finding out what the requirements for truth of the truth-evaluable representation are. These are complex, as the overall structure involves multiple referential representations whose referents have to stand in certain relations to each other for the representation to be true.

The overall structure is that of a conditional with a conjunction in the antecedent. So the representation will be true, if either <Tobi,Conny> is not a referent of ⌜taller than⌝ , or if <Conny, Laila> is not a referent of ⌜taller than⌝, or if all three pairs, <Tobi, Conny>, <Conny, Laila>, and <Tobi, Laila> are referents of their respective token representations of ⌜taller than⌝.[131]

[131] As before one would, to be quite precise, have to distinguish between each token representation exemplified by the words “taller than”. However, since it was (reasonably) assumed that they are of the same type, there is no harm done in considering them together and thinking of all the pairs as being referents of the same type.

So there is a number of different relations between the referents of the representations that suffice for the truth of the truth-evaluable representation.

To find out whether one of these relations between the referents are guaranteed to hold in virtue of how the reference of the constituent representations is determined, the reference realizing properties determined by the referential representations must be clarified.

The reference realizing properties of the pair's singular referential representations, ⌜Tobi, Conny⌝, ⌜Conny, Laila⌝, and ⌜Tobi, Laila⌝, need not be given in detail here. It will be enough to note two uncontroversial features. First, the two uses of each name are token representations of the same type and they therefore determine the same reference realizing properties. Second, the reference realizing properties of the pairs suitably depend on those of the singular referential representations referring to the elements of the pairs.

The more interesting reference realizing properties for present purposes are those of the three uses of ⌜taller than⌝. It can be assumed that they are all representations of the same type and that they thus determine the same reference realizing properties. A theory of reference for the relation, however, is not so easy to find: A simple description theory is at least not obviously forthcoming, and it is also slightly unclear how a causal theory may work in the present case. So I will have to abstain from speculating about the theory of reference. In either case, however, we arguably do know at least some things about the reference realizing properties determined by ⌜taller than⌝, and so I will suggest some plausible reference conditions, which are roughly in

line with linguistic analyses of comparatives.[132] Whether the resulting reference realizing properties really are the correct ones, however, will remain conditional on a true theory of reference actually yielding reference conditions along these lines. Nonetheless, the exercise will be instructive. Here is the suggestion for the reference conditions of ⌜taller than⌝:

An ordered pair of objects is the referent of ⌜taller than⌝, iff the first member of the pair has a size greater than the second member of the pair.

These reference conditions invoke the property of size as well as assuming that these sizes are comparable as smaller and greater. These assumptions should not be objectionable, as there is no intention to give a reductive account of the relation eliminating the comparative relations. These reference conditions, together with the assumptions about the singular representations, imply the necessary truth of our target representation in the following way:

Let's assume that ⌜taller than⌝ does refer to the pair <Tobi, Conny> and to the pair <Conny, Laila>. If this is the case, we know from the reference realizing properties of ⌜taller than⌝ that Tobi has a size greater than Conny, and that Conny has a size greater than Laila. Next we ask what would have to be the case for ⌜taller than⌝ to have the pair <Tobi, Laila> as referent. This pair is a referent of ⌜taller than⌝, if and only if Tobi has a size greater than Laila.

---

[132] In linguistics it is common to understand the relation 'taller than' as a relation between sets of degrees of size. See for example Kennedy (2014) for an overview over the semantics of comparatives, as well as Klein (1980).

Now if the pair <Tobi, Laila> is bound to have these reference realizing properties given that the pairs <Tobi, Conny> and <Conny, Laila> are referents of ⌜taller than⌝, then the requirements for truth are guaranteed to be satisfied, because they are satisfied if all three pairs are referents of the ⌜taller than⌝ relation. That Tobi indeed has a size greater than Laila, and that therefore the pair <Tobi, Laila> is a referent of ⌜taller than⌝, if <Tobi, Conny> and <Conny, Laila> are, follows from the facts, derived from the reference realizing properties that Tobi has a size greater than Conny and that Conny has a size greater than Laila. So it is the case that the pair <Tobi, Laila> is a referent of ⌜taller than⌝, if the pairs <Tobi, Conny>, and <Conny, Laila> are.

This is just what needs to be established to explain the necessary truth of the truth-evaluable representation. For the requirements for truth are bound to be satisfied in virtue of the relationship between the reference realizing properties of the referential representations.

The examples discussed here, ranging from simple conceptual and logical truths to more complicated relations, demonstrate that the present account, at least in principle, has the resources to account for the necessary truth of these kinds of representations.

This, however, is still not very surprising, as it would have been expected of an account developed from an account of analyticity. In the next sections, I suggest how the theory may explain the necessary truth of representations involving singular referential representations. These broadly essentialist truths have a special feature which has been noted and discussed in the literature: They seem to require for their truth that the object represented by the singular referential representation exists. So if such

representations are necessarily true, they seem to imply the necessary existence of the object represented by the singular representation. This is a puzzle that needs to be solved, if one is unwilling to accept that everything exists necessarily. Connected to this puzzle about necessary existence are questions about merely possible objects, objects that do not actually exist, but may exist, as well as impossible objects, objects that just cannot exist. It is the topic of the following sections to clarify how these puzzles may be dealt with in the present framework.

Interlude: Necessary Existence and Mere Possibilia

In this section I will argue that puzzles surrounding necessary existence, mere possibilia, and necessarily non-existent objects can be dealt with in the present framework, by distinguishing between different readings of sentences involving names.

Puzzles arise once a theory of necessity allows for (broadly) essentialist truths, truths that attribute an essential property to an object. A prominent example is the representation expressed by the sentence "Socrates is human" which will in the following be assumed to be necessarily true.[133] The puzzle arises, because it seems that if it is necessarily true that Socrates is human, then Socrates exists necessarily, for how can it be true at all that Socrates is human without Socrates existing?

The solution, I argue, is to distinguish between two readings of essence attributions. On one reading of such sentences, it is required for the truth of the representation expressed that the singular representation expressed by the name has a referent. On

[133] I do not presently take a stand on which specific essence attributions are true. But some will be true, and so the skeptic can substitute her favorite example in place of the one used here.

this reading, no such representation turns out to be necessarily true according to the present theory. On another reading, however, it is not required for the truth of the representation that the singular representation has a referent. On this reading the representation may turn out to be necessarily true according to the theory, but it no longer implies the necessary existence of a referent of the singular representation. This diagnosis is consistent with standard solutions to the puzzle of necessary existence.

First I will discuss the puzzle of necessary existence and argue for the suggested solution. These findings are then shown to impact discussions about merely possible and impossible objects.

Necessary existence

Some truths involving singular representations seem to require for their truth that the object to which the singular representation refers, exists. If these representations can be necessarily true, a puzzle arises, for it seems that we can straightforwardly deduce the necessary existence of the objects referred to by the singular representation. Here is how Bob Hale puts the argument: "It is in the nature of Aristotle to be a man—being a man is part of what it is to be Aristotle. […] So it is necessary that Aristotle is a man. But nothing can be true of Aristotle unless he exists. So it cannot be true that Aristotle is a man unless he exists. So Aristotle necessarily exists!"[134] This is problematic, since ordinarily we would think that people and ordinary objects could have failed to exist: If Aristotle's parents would not have met, he would never have been born, and so would not have existed. Similar stories can be told for all ordinary objects. It even seems true that nothing

[134] Hale (2013) p. 211.

could have existed. So unless we are prepared, like some philosophers, to believe that everything exists necessarily, we have to reject some part of the argument for necessary existence.[135]

First, I will discuss a standard solution to the puzzle, then move on to Kit Fine's criticism of the standard solution and his own proposal, and finally look at Bob Hale's recent take on the issue. Finally I argue for my own solution, which flows from the representation-framework, and show how it unifies these other proposals.

The argument for necessary existence and the standard solution

The argument for the necessary existence of an ordinary object can be spelled out in argument form as follows. I use Kit Fine's formulation here, as it brings out nicely the paradoxical character of the puzzle:

"(1) It is necessary that Socrates is a man;

(2) It is possible that Socrates does not exist;

(3) Therefore it is possible that Socrates is a man and does not exist"[136]

Premiss (2) is clearly true, premiss (1) may be controversial, but can be substituted by any other property Socrates necessarily has, for example his self-identity, for the argument to go trough. The conclusion (3) however, seems unacceptable, "for how can Socrates be a man without existing?"[137] And so we might be led to conclude that (2) is false after all, and that it is not possible that Socrates does not exist.

---

[135] Williamson (2013) claims that everything exists necessarily.

[136] Fine (2005) p. 328

[137] Fine (2005) p. 329

There is a more or less standard solution to the problem, as Fine points out, which distinguishes between a qualified and an unqualified reading of the modality in question. "Under the 'unqualified' reading, a proposition concerning certain objects will be necessary if it is true in every world,[…]. Under the 'qualified' reading a proposition concerning certain objects will be necessary if it is true in any world in which those objects exist, […]"[138] So according to the standard solution, the mistake is a fallacy of equivocation: drawing an unqualified conclusion from the qualified premiss (1).

This comes out more clearly in possible-world talk. Premiss (1), rephrased in terms of possible worlds says, on the standard solution, that in all possible worlds, if Socrates exists, then he is a man. Premiss (2) says that there is a possible world in which Socrates does not exist. All that follows from this is that there is a possible world in which Socrates does not exist and in that possible world, if Socrates exists, then he is human. But it does not follow that there is a possible world in which Socrates does not exist and in which he is human.

Fine, however, has some quarrel with this standard solution. His main objection is that there is no clear linguistic motivation for diagnosing an equivocation between the qualified and unqualified sense of 'necessarily', for we do not seem to have any evidence that there are different senses at issue. So, conditional on the quality of Fine's arguments, it seems that, absent a good independent reason to treat the two readings of 'necessarily' as distinct, it is at least doubtful that the diagnosis offered by proponents of the standard solution is correct.

---

[138] Fine (2005) p. 330

Fine's solution: worldly and unworldly predicates

To solve the puzzle, Fine draws a distinction between 'worldly' and 'unworldly' predicates and the resulting sentences. The idea is that if a true sentence is formed with a worldly predicate, then it is existence-entailing. If it is formed with an unworldly predicate, then it is not existence entailing. The distinction between worldly and unworldly predicates is motivated by an analogy to a distinction within the temporal domain. Here a distinction can be drawn between tensed and tenseless expressions. A tensed expression is, to use Fine's example, 'exists', for "Socrates exists" is true at a time where Socrates exists, and false at times where Socrates is no more or was not yet. Opposed to that are tenseless expressions, such as 'is a man'. For "Socrates is a man" is true, and "cannot properly be said to be true or false at a time".[139] Corresponding to this distinction, a distinction can be made between sempiternal truths and eternal truths, where sempiternal truths are tensed sentences that are always true and eternal truths are tenseless sentences that are true simpliciter. Fine notes that the distinction between eternal and sempiternal truths can be obscured, if we admit an extended sense of truth-at-a-time, for in that extended sense, eternal truths can be said to always be true. So sempiternal as well as eternal truths will just be true at any given time. But he insists that even so, a distinction can be drawn, for "a tensed sentence will be true at a time *because of how things are at that time*, while […] a tenseless sentence will be true at a time *regardless of how things are at that time*."[140]

---

[139] Fine (2005) p. 322
[140] Fine (2005) p.323. Emphasis mine

If we accept this distinction, Fine argues, we should accept an analogous distinction in the modal sphere. Here "one can draw a distinction between worldly and unworldly sentences according to whether they can be properly said to be true or false in a world."[141] Corresponding to this distinction is a distinction between transcendental and necessary truths, where the transcendental truths will be true regardless of the circumstance, and the necessary truths are true whatever the circumstance.

This distinction solves the puzzle under the assumption that 'man' is unworldly, while 'exists' is worldly, because if we accept that 'man' is an unworldly predicate, then it doesn't follow that the truth of "Socrates is a man" requires the existence of Socrates. And so the conclusion can be accepted.[142]

Fine's solution seems to have an advantage over the standard solution in terms of different senses of the word 'necessarily', because no equivocation and thus no mistake on the side of the reader of the argument must be diagnosed. Rather, an independently plausible distinction between truths that require the existence of the object they are about and those that don't is invoked to explain the acceptability of the conclusion.

One must be slightly careful in evaluating the dialectic advantage of this position, however. For even though Fine can give both premises an unqualified reading, and thus pay tribute to our intuitions when reading the argument, he has to attribute to the reader of the argument, who is puzzled by it, a different mistake,

---

[141] Fine (2005) p 324

[142] Fine (2005) also distinguishes between three different senses of 'necessarily' depending on whether a proposition is transcendental, necessary, or a mixture of both. This is important for the readings of the modal idiom in the argument leading to the puzzling conclusion. For the current exposition of the argument, however, no detailed discussion of this is necessary.

for the reader has a mistaken reading of the conclusion: he fails to notice that the predicate 'human' is unworldly, and thus a sentence formed with it and a singular term may be true even though the object the singular term purports to refer to, does not exist.

The delicate linguistic dialectic, however, need not concern us too much. For Fine's solution, on second thought, is not so different from the standard solution after all. For notice that it is not just the case that transcendental truths do not require the existence of the object they are about for their truth, it should also be the case that if the object exists, it has the property indicated by the predicate. But if this is so, we have a quite simple and more perspicuous formulation for the transcendental truth that Socrates is human: If Socrates exists, Socrates is human. This is a conditional truth, conditional on the existence of Socrates. But if we substitute this way of formulating the transcendental truth for "Socrates is human" in the argument above, we just arrive at the standard solution.

The argument between Fine and the proponents of the standard solution should then not be thought to be one that concerns the substance of the solution: Both solutions work by conditionalizing the truth in the scope of the necessity-operator (Socrates is human), explicitly or implicitly, on the existence of the object it is about. The quarrel should be thought of as an argument where to locate the reasons for giving it such a reading, given that the surface form is not conditional. The standard solution locates the reasons for the conditional reading in an allegedly ambiguous necessity-operator, while Fine locates the reasons for giving the sentence a conditional reading in the predicate we use to form the sentence.

Hale's solution: restricting quantifier-elimination

Bob Hale's solution is slightly different still. It rests on the insight that in the modal domain we need to take care what we quantify over, and suggests restricting the elimination rule for the universal quantifier: The universal quantifier should only be eliminable in contexts where the term that is eliminated for has a referent.[143] Essence attributions, on his view, should be possible, without this attribution entailing the existence of the object that has this essence. This observation seems licensed by the thought that in specifying what it is for a thing to be the thing it is, we need not at the same time make any commitment to the existence of the thing we are specifying the essence of.

How does this solve Fine's puzzle? Hale thinks that we should reject premise (1) as it stands, for it does not follow from its being the essence of Socrates to be a man, that Socrates is necessarily a man. For, as Hale puts it, "we should not accept (1) unless we are prepared to insist that there can be such a thing as *what it is to be [Socrates]* only if there is such a thing as [Socrates]. But it is not, in general, at all plausible to hold that there is such a thing as what it is to be X only if X exists, i.e. only if there is such a thing as X."[144] Rather, what follows from its being part of the essence of Socrates that he is a man, can be expressed more perspicuously by the following conditional attribution: Necessarily, for all x (If x=Socrates, then x is a man)[145]. This conditional statement

---

[143] Hale (2013) p. 209ff.

[144] Hale (2013) p. 215. Hale uses Aristotle, instead of Socrates as an example. To stick with Socrates, the example Fine gives, all occurrences of 'Aristotle' in the quote have been replaced by 'Socrates'. It obviously doesn't make a difference to the basic point.

[145] Hale (2013) p. 216

doesn't allow to straightforwardly deduce the undesirable conclusion that possibly, Socrates both is a man and doesn't exist. For this to follow, it would have to somehow follow from the conditional that necessarily, Socrates is a man. This may be thought to be effected by an instance of the universal generalization, namely by taking x to be Socrates. In this case, it would follow that necessarily, if Socrates = Socrates, then Socrates is a man. Since Socrates=Socrates is necessary, it follows, that necessarily, Socrates is a man. To avoid this conclusion Hale suggests that we have to restrict the elimination rule for the universal quantifier, for by introducing the instance x=Socrates, we "illicitly smuggle in" the assumption that Socrates exists.[146] We should instead replace the standard rule by the following rule of free universal quantifier elimination: "From $\forall x\, A(x)$, together with an existence-entailing premise $F(t)$, we may infer $A(t)$, where t can be any term."[147] The crucial difference is the need for an existence-entailing premise for the instantiation to be allowed, and if we want to eliminate the quantifier in the modal context, we need a necessary existence entailing premise. But since no premise that entails that Socrates necessarily exists is true, the instantiation is not allowed. So the puzzle is solved by moving to a free logic.

This solution, while taking a detour through the theory of quantifier elimination rules of a free logic, arrives at a conditional reading of premise one as well: It is not unconditionally necessarily true that Socrates is a man, but it is necessarily true that if Socrates exists, then he is a man. So the substance of the solution is the same as that of the standard solution, but instead of

---

[146] Hale (2013) p. 209
[147] Hale (2013) p. 209

making the conditional character of the necessary truth explicit, it puts the conditional in the quantifier elimination rules. So one may either make the conditional reading explicit, as in the standard solution, one may make it almost explicit via not requiring the existence of the object for the truth of the proposition, or one may make it implicit in the quantifier elimination rules. Which way we go does not change the substance of the solution. The difference lies in the diagnosis of what gives rise to the truth not depending on the existence of the object referred to by the name: By amending the quantifier elimination rules, we can stick to the surface subject-predicate form, and by making the conditional explicit, we can keep the quantifiers, but have to amend the subject-predicate expression to be read as a conditional. Either way, it is not true unconditionally that necessarily Socrates is a man, it is merely necessarily true that if he exists, he is a man.

Necessary existence in the current framework

The essentialist truth expressed by the sentence "Socrates is human" is necessarily true, according to the present theory, if the reference realizing properties of the singular representation ⌜Socrates⌝ and the general representation ⌜human⌝ are such that they guarantee the satisfaction of the requirements for truth. This can only be the case, I will argue, if the requirements for truth do not require that the singular representation ⌜Socrates⌝ has a referent. If the requirements for truth require the existence of Socrates, the representation is not necessarily true. This is because nothing can guarantee that anything has the reference realizing properties of Socrates. Consequently, the satisfaction of

requirements for truth which require the existence of a referent cannot be guaranteed by reference realizing properties of constituent representations. And so no representation requiring such a thing can be necessarily true.

To make this more explicit, I will go through the example used above, the necessarily true representation expressed by the sentence "Socrates is human".[148] For the representation expressed by "Socrates is human" to be necessarily true, the reference realizing properties determined by ⌜Socrates⌝ and ⌜human⌝ must be related such that the requirements for truth are guaranteed to be satisfied. Whatever else the reference realizing properties determined by the representation ⌜Socrates⌝ may include, let us for present purposes assume that they do include the property of being human. Thus, something is a referent of ⌜Socrates⌝, only if it is human.

Correspondingly, we may assume that the reference realizing property of ⌜human⌝ is given by the trivial reference realizing property of being human. Thus something is a referent of ⌜human⌝, if and only if x is human.

The sentence "Socrates is human" has a subject-predicate structure and so it may seem that the requirements for truth are analogous to representations expressed by other sentences with this structure, like "bachelors are unmarried". So a first shot at the requirements for truth for the representation expressed by "Socrates is human" are as follows:

---

[148] Again it does not matter whether this particular essentialist truth is true, as any other example could be used to make the point.

⌜Socrates is human⌝ is true, iff the referent of ⌜Socrates⌝ is a referent of ⌜human⌝.

This statement of the requirements for truth, however, is not clear on whether there is no referent of ⌜Socrates⌝. Should the requirements count as satisfied, if there is no referent of ⌜Socrates⌝? This would be analogous to sentences with subject-predicate structure expressing a representation involving only general representations like "bachelors are unmarried". Or should they count as not satisfied, if there is no referent of ⌜Socrates⌝? This latter option seems to cohere better with the common understanding of subject-predicate sentences involving names as entailing the existence of the object named, if they represent truly.

By reformulating the right hand side of the biconditional, we can make these readings explicit. Each option corresponds to a different reading of the sentence expressing the representation.

i) If something is a referent of ⌜Socrates⌝, then it is a referent of ⌜human⌝.

ii) There is a referent of ⌜Socrates⌝, and it is a referent of ⌜human⌝.

While the first reading does not require the existence of Socrates, for it is true if there is no referent of ⌜Socrates⌝, the second reading does require the existence of Socrates, for it is only true if there is a referent of ⌜Socrates⌝.

Interestingly, only the first reading, the one not requiring ⌜Socrates⌝ to have a referent, turns out to be necessarily true, given the reference realizing properties assumed for both

⌜Socrates⌝, and ⌜human⌝. For if the reference realizing properties are as given, then, if there is a referent of ⌜Socrates⌝, it is also a referent of ⌜human⌝. So the relationship between the reference realizing properties guarantees the satisfaction of the requirements for truth, if the requirements for truth are given as in i) above.

The requirements for truth as given in ii) are not guaranteed to be satisfied by the reference realizing properties, however, because there is no way to guarantee that there is an object with the relevant reference realizing properties.

So the only reading on which it is true that the representation expressed by "Socrates is human" is necessarily true, is if we understand the sentence as expressing a representation with conditional requirements for truth along the lines of i). This result is parallel to both the standard solution and to Fine's solution of the puzzle of necessary existence, because it is the mark of transcendental truths, that they do not require the object they are about to exist to be true.

This result, however, was not arrived at by considering linguistic evidence purporting to identify either an ambiguity in the word 'necessarily' or by distinguishing between different kinds of predicates, but by showing that only the conditional reading provides us with a necessary truth. This is consistent with both the standard solution, and Fine's solution, as they may both be viewed as different theses about how our language succeeds in expressing two different representations with the same sentence, each with their own requirements for truth. The present theory can stay neutral on linguistic issues, as long as some conditional reading is available at all. What the theory is committed to, is that

only if the requirements for truth do not require the existence of a referent, can there be a necessary truth involving singular representations. This also implies that no representation representing something to be a necessary existent will turn out true on the present theory. Whether this is a feature or a bug of the theory, will be discussed in the section on mathematical truths.
In the following examples of necessary truths involving singular representations, I will assume that the requirements for truth are conditional analogously to the ones outlined in i).

Mere possibilia
Just as there are contingently existing objects, there seem to be, at least as a manner of speaking, contingently non-existing objects. These objects do not exist, but they might have existed.[149] Examples are the sister I could have had, the planet Vulcan, the possible fat man in the doorway, and so on. What is puzzling about such mere possibilia, is that there seems to be a sense in which they do exist, for we seem to talk as if they do. But, of course, merely possible 'objects' do not exist, because they exist merely possibly.[150] So what are these mere possibilia, and how does the current theory account for them?
First, it is to be noticed that the basic considerations about reference hold even if no object satisfies the reference conditions

[149] Talk of possible objects is already quite misleading, for it seems to imply that there actually are objects which are merely possible. This should not be assumed from the outset, of course, and as I will argue, should be treated as a manner of speaking. I will, however, follow common usage and talk of possible objects here, and sometimes put the term 'object' in scare-quotes to highlight that it should not be read as implying that there really are objects which happen to be merely possible.
[150] Since actual objects are also possible objects in virtue of being actual, I follow common usage and talk of *merely* possible objects to talk about non-actual objects.

for a referential representation. The reference determining properties for the representation ⌜Vulcan⌝, for example, presumably determine some reference realizing properties that are just not had by anything.[151] Still, they could have been had by some object, for Le Verrier could have been correct that there is a planet which explains the changes in the perihel of Mercury. So we may conclude that a representation represents a merely possible object, so to speak, if nothing has the reference realizing properties determined by the representation, but it is not impossible that something has these reference realizing properties.

This observation helps to think of merely possible 'objects' as not quite so puzzling. For we represent possible objects just like actual objects, we just sometimes fail to represent anything that exists with our representations, though we might. What distinguishes possible from actual objects is, unsurprisingly, that the possible objects do not exist, but may have existed.

Still it may seem puzzling how we can quantify over all these objects that do not exist. We seem to be saying on the one hand that there are possible objects, but on the other hand that these possible objects are not the referents of our representations. To see how this puzzle may be solved, let us consider in slightly more detail the example of the possible object Vulcan.

Vulcan is a merely possible 'object', Vulcan does not exist, but it could have existed. So the reference determining properties of the representation ⌜Vulcan⌝ determine some reference realizing

---

[151] There may be some doubts that singular terms without a referent can have determinate reference conditions. This potential worry will be discussed later. Here I merely note that we should be able to distinguish between names that purport to refer to an object and those that just fail to be representations at all.

properties, which are not had by anything. Still, and this is what makes it puzzling, we speak as if something may have the relevant reference realizing properties, for we speak of the possible object Vulcan, and this possible object seems to have them.

The first thing to note about Vulcan is that not all true representations involving ⌜Vulcan⌝ require for their truth that there is an object with the relevant reference realizing properties. Just as in the case of Socrates' being human, the truth of the representation expressed by "Vulcan is a planet" need not require the existence of Vulcan for its truth, it may just require it to be the case that if there is something that has the reference realizing properties of ⌜Vulcan⌝, then it is a planet. So there is no need to worry about having to accept the existence of mere possibilia to make some representations involving reference to non-actual objects true. It is for the same reason that we need not worry about the necessary existence of Socrates, even though we attribute a property to him he necessarily has. The truth of these representations just does not require the existence of the object they are about.

While this much may be granted, explicit reference to mere possibilia, as in "there is a possible planet between Mercury and the sun" or as in "Vulcan is a possible planet" may remain puzzling. For here we seem to be saying explicitly of something that merely possibly exist, that it exists. To dispel this puzzle, we can once more distinguish between two representations that may be expressed by these sentences. The first requires for its truth the existence of a possible planet between Mercury and the sun, and is therefore false. The second does not require for its truth that

there is a possible planet between Mercury and the sun, but merely that it is not necessarily not the case that there is a planet between Mercury and the sun.

So sentences apparently involving reference to possible objects can be understood either as representations which require for their truth the existence of merely possible objects or they can be understood as representations only requiring that it is not impossible that the singular representation apparently referring to the possible object has a referent. This can be made more explicit by reformulating the sentences above as "Possibly, there is a planet between Mercury and the sun" and as "Possibly, Vulcan is a planet".

Is it plausible to read quantification of possibilia in this way? Is this really what we are representing, when we speak of possible objects? The reason for thinking that this really is the most charitable interpretation of our possibilist discourse comes from considering what it is to be a *merely* possible object: It is partly to be an object that does not exist.[152] So if the truth of the representations expressed by the sentences above would require the existence of a possible object, then they turn out to be false, for possible objects do not exist. The reading which does not require the existence of possible objects, however, can interpret the sentences as expressing a truth, for it is clearly not necessarily not the case that there could not have been anything with the relevant reference realizing properties.

---

[152] This may be disputed by philosophers who make a distinction between different ways of existing. Williamson (2013), for example, argues that we should distinguish between existing concretely and abstractly, and possible objects that exist abstractly. But then Williamson also thinks that everything exists necessarily, a thesis which can be safely denied on the present theory.

So the most charitable interpretation, given that possible objects do not exist, is the interpretation on which their existence is not required for the truth of the representations expressed by the sentence. This case is further bolstered by considering impossible 'objects', 'objects' which could not have existed in the next section, for sentences involving apparent reference to them can be treated analogously to sentences involving apparent reference to mere possibilia.

Much of the philosophical debate around possibilia is conducted within a possible-worlds framework. Here the debate is often framed as a debate between actualists, philosophers who deny the existence of possibilia, and possibilists, who take possibilia to exist. Given that both parties to the debate accept a possible world analysis of necessity and possibilia, the actualists are committed to finding actualistically acceptable substitutes for possible objects. Since possible worlds are one special kind of possible object, actualists in this debate are usually ersatzists, trying to find substitutes in the actual world to go proxy for merely possible objects. Possibilists have commonly had the upper hand in this debate, as there are commonly severe problems with the suggested proxy-reductions.[153] The difficulties the proxy-actualists have in finding adequate substitutes for possible objects should not be surprising given the poor explanatory prospects of reductive ersatzism outlined in the first part.

The present position also denies the existence of possibilia, but it does not suffer from the defects of standard proxy-actualist positions, because it is not committed to a possible-worlds analysis of necessity. Since there is an explanation of necessity

[153] For an overview of actualist and possibilist positions see Menzel (2016), for further criticism of the standard actualist proxy-reductions see Fine (2005)

(and accordingly possibility) without possible worlds, we can use this explanation of necessity to understand our quantification over possible worlds and possibilia generally.

Here is how Kit Fine, who suggests a similar reformulation in terms of necessity and possibility, puts the point: "Many philosophers seem to have followed Lewis in supposing that they must either go with *proxy* reduction ('ersatzism') or accept possible worlds realism. But this is a false dilemma. For as I have indicated in previous work, it is possible to provide a straightforward non-proxy reduction of possibilist discourse. The basic idea is to take modality as primitive and to treat the possibilist quantifier 'there is a possible object *x*' as equivalent to 'possibly there is an object *x*'[…]"[154] The suggested translation, Fine admits, is not always quite as straightforward, but the basic idea is the same that was outlined here: Mere possibilia are to be understood in terms of possibility, and not vice versa.

To sum up, reference to mere possibilia need not be taken to be a miracle. It can be understood like reference to everything else. It just sometimes so happens that a reference relation fails to hold, because there is nothing with the appropriate reference realizing properties, but there could have been. This is as it should be, for actual objects are also possible objects, and so their treatment should not differ markedly from merely possible objects. The difference is just that in the case of merely possible 'objects' nothing has the relevant reference realizing properties.

---

[154] Fine (2005) p. 224-225.

Impossible objects

An impossible object, I will argue, is best understood in the same vein as a possible object, but as one which does not possibly exist. So it will be false to say of an impossible object that it possibly exists. Nonetheless there seem to be some truths about impossible objects, for example that they cannot exist. This may seem puzzling, but the understanding of possible objects from above helps to dissolve this puzzle.

A classic example of an impossible object is a square circle. It is commonly taken to be impossible that anything can be both square and a circle, consequently the square circle (necessarily) does not exist. Nonetheless it seems true that it is in the nature of the square circle to be both square and a circle. How can all this be the case? The first step is to give talk about impossible objects an appropriate reading. Just as merely possible objects, they do not exist, and so are not objects at all. Possible objects could have been objects, but are not, while impossible objects lack the possibility of having been objects. We should take quantification over impossible objects to be a manner of speaking, just as talk of mere possibilia is a manner of speaking. Both are not objects, because they do not exist, it is just that there are some modal truths involving representations that purport to have them as referents. In the case of possible objects, it is possible that something with the characteristics of the possible object exists, and in the case of impossible objects it is not possible that something with the characteristics of impossibles object exists.

Let me show how this works in the present framework. It is true that there are no square circles, if there is no referent of the representation ⌜square circle⌝, and this is necessarily true, if something about the reference realizing properties precludes that

anything has the reference realizing properties of ⌜square circle⌝. Let us assume that the reference realizing properties determined by ⌜square circle⌝ are that something which is a square circle, is both square and a circle. Since everything which is a square is not a circle, these properties cannot be had by anything, and so it is necessarily true that nothing is both square and a circle.

Even though necessarily, nothing is a square circle, there are some true representations about square circles. One of them is the truth that a square circle is both square and a circle. This seemingly paradoxical result can be explained with the strategy outlined for possibilia. Consider first the requirements for truth of the representation expressed by the sentence "the square circle is square and a circle", once more conceived of as not requiring the existence of the object referred to by the representation in subject position for their truth.

⌜the square circle is square and a circle⌝ is true, iff if there is a referent of ⌜square circle⌝ , then it is a referent of ⌜square⌝ and a referent of ⌜circle⌝.

These requirements are guaranteed to be satisfied in virtue of the reference realizing properties of ⌜square circle⌝, for we can assume them to be the following: Something is the referent of ⌜square circle⌝, iff x is square and x is a circle.

The treatment of impossible objects is analogous to the treatment of possible and actual objects. What is special about them is that the representations attempting to represent such objects determine reference realizing properties which cannot be had by anything. Still, that these reference realizing properties are determined as they are, gives rise to necessary truths about square circles, for the

requirements for truth of the appropriate representations are guaranteed to be satisfied by the reference realizing properties.
I argued that quantification over possible objects as well as over impossible objects should be treated in essentially the same way and that talk about possibilia and impossibilia can be taken to state a regular modal truth which asserts the possibility or impossibility of the existence of some object. This renders talk of mere possibilia and impossibilia unmysterious: It is like talk of any other object, just that either the object we purport to refer to happens to not exist, or cannot exist. This explains why there can be truths involving non-existent objects, and why it is still true that these objects do not or even cannot exist.

Reference realizing properties for non-existent objects?
In the discussion it was assumed that the reference determining properties of singular representations which purport to refer to something, but where this something fails to exist, do determine reference realizing properties, which are not had by anything.
This may be doubted, for it may be thought that in order to get determinate reference realizing properties, at least for singular representations, an object has to exist. For how are there enough reference determining properties of the representation to ever determine uniquely identifying reference realizing properties, if the object does not exist? In the case of existing objects it may at least somehow be explained how a causal theory of reference succeeds in determining a unique referent, but how can this work, if there is nothing there to which the name can be given? In other words, how do we succeed in singularly referring to non-existent objects?

On one reading this problem just does not arise. For we do not, and cannot, refer to nonexistent objects. This is because a reference relation can only hold between a representation and an object, if the object exists. And non-existent objects do not exist. So there is no problem about referring to non-existent entities, since no referential representation can refer to non-existent entities.

Still, we do sometimes use referential representations which happen not to refer to anything in (sometimes true) truth-evaluable representations. For this to be the case, the reference determining properties must be such that they determine some reference realizing properties, which are then not had by anything. To clarify these remarks, let me give an example:

Suppose that Leverrier introduced the name Vulcan thus: "I call the planet between the sun and Mercury, which explains the perturbations of Mercury's perihel by the name 'Vulcan'." With this reference fixing description, he intended to fix the reference of 'Vulcan' to a unique planet between the sun and Mercury. If the naming had succeeded, that is, if there had been such a planet, then the name 'Vulcan' would have referred to it. Also, most of the properties used to pick out the referent would, in this case, not have been essential to being Vulcan, for Vulcan may, after having been so named, been hit by something and thrown to a different orbit, such that it was no longer between the sun and Mercury. Also it may have turned out that this planet, even though it exists, did not in fact explain any perturbations in Mercury's orbit. Still Leverrier may have succeeded in giving a name to it, which henceforth represents that planet, and therefore determines some

reference realizing properties. In the simplest case, this would have been the property that it is the same planet as Vulcan.[155]

The trouble is that there simply is no such planet, and we can appeal to no planet to which the object referred to, is identical when giving the reference realizing properties. So what may the reference realizing properties of 'Vulcan' be?

A partial, but slightly unsatisfying answer is to assume that the reference determining properties simply do not suffice to specify the reference realizing properties which would pick out an object and so the name could not refer to any object. So one may be lead to the conclusion that it is not even possible that Vulcan exists, because there is no way to refer singularly to Vulcan.[156]

This problem and the negative conclusion is discussed by Kripke for non-actual natural kinds which he also takes to be subject to a causal account of reference:[157] Kripke asks us to consider Unicorns. Unicorns do not exist, they are a mythical species. So *prima facie* it is unclear how reference to them is fixed. Contrast this with existing species like Tigers. Tigers are not defined by their superficial properties, but by their actual internal structure, which may not be known to the people who introduced the word 'tiger' by pointing to a specimen of the species tiger. Unicorns, in contrast, do not exist and so there is no actual unique unicorn-species with a certain internal constitution, which could be

---

[155] What it takes to be the same planet is of course left open here, and in need of further clarification. Some general suggestions were made in the semantic preliminaries.

[156] This result seems to be quite general, if we allow the objects that actually exist a role in reference determination, as is arguably the case in a causal theory of reference. For the role in reference determination will not be filled by anything if the object one attempts to give a name to, does not exist and so reference for a unique object will not be determined.

[157] Kripke (1980) p. 23/24 and 156-158.

named. Kripke writes: “regarding the several distinct hypothetical species, with different internal structures (some reptilic, some mammalian, some amphibious), which would have the external appearance postulated to hold of unicorns, one cannot say which of these distinct mythical species would have *been* the unicorns.”[158] No distinct species could be picked out with the word ‘unicorn’, and thus nothing could possibly be a unicorn.

Note that at no point is Kripke denying that something could have turned out to be denoted by the word ‘unicorn’. However, as it stands, no such object was indicated, and so it does not suffice to represent a species, for it is lacking the connection to an existing specimen of the species necessary for unique reference-fixing.[159]

The same, so one may argue, goes for particulars. Here is how Barcan Marcus makes a similar point: “In summary, modal discourse need not and should not admit possibilia despite the elegance of the generalization. Dispensing with possibilia is grounded not in the unavailability of criteria of identification but rather in the fact that identity is a relation for objects already given. Putative possibilia are not fleshed out with that complement of properties, relations, and a locus in the actual

---

[158] Kripke (1980) p. 156-157.

[159] The point may be made more precise in a two-dimensional semantic framework, as it is used, for example, to make sense of indexicals: If uttered at a different possible world, the word would have a certain reference fixed, which is then rigid over all possible worlds, that is, if we consider another world as actual, the word would have had a referent. However, given that no reference is fixed in the actual world, the word does not refer to anything in every possible world, if considered as counterfactual. The applicability of the framework, however, depends on a division of the reference determining properties into those we hold fixed while considering other worlds as actual, and those which vary across those worlds. This division is not available, if we move from words to representations, for representations are individuated by the reference realizing properties they determine.

order—or, if material, in the physical order—that would enable them to count as objects at all."[160]

Treating reference to possibilia in the way Kripke and Barcan Marcus do, and to deny that there could be a referent to (singular) representations representing something that does not exist on the grounds that no determinate reference is achieved, if no referent exists, seems to leave us with a host of new puzzles, however. In the representation-framework, the suggestion would amount to the claim that the reference determining properties of representations do not suffice to determine reference realizing properties. But if so, how can we say that such representations are even meaningful representations? And if they do not determine reference realizing properties, how do they give rise to apparent truths expressed by sentences like "unicorns do not exist"? A sentence expressing a necessary truth on Kripke's account.

Nonetheless there is an important insight in Kripke's and Barcan Marcus' remarks. We cannot easily refer to objects which merely possibly exist. So how can the insights of Barcan Marcus and Kripke be accommodated without having to suppose that names like 'Vulcan' could not have referred to anything at all?

There may be a way to accommodate these insights without having to abandon representing mere possibilia. For even if no referent is picked out at an attempted baptism, it need not be assumed that no reference realizing properties are determined at all. After all, in the spirit of the current framework and the discussion of possibilia above, it can be expected that some reference realizing properties are determined which just happen to not be had by anything. For we do want to make a distinction

---

[160] Marcus (1995), p. 213.

between names or other alleged representations that fail to have certain representational properties, and those that do have representational properties, but which fail to have a referent, that is, a distinction between objects which fail to be fully representational, and representations which do not have a referent.[161]

If there is no referent, the reference realizing properties cannot rely on the existence of an object referred to, still there is something that can be said of them. Take again Vulcan as an example: Since it was introduced to explain the changing perihel of Mercury, we may assume that its reference realizing properties are something like this: being the unique planet between the sun and Mercury which explains the changes in the perihel of Mercury. Since no planet actually is such that it explains this, Vulcan does not exist. Still, it could have been the case that there is such a unique planet, if a planet had explained the changes in the perihel of Mercury. Given that the reference realizing properties are as suggested, it is necessarily true that if Vulcan exists, then it explains the changes in the perihel of Mercury.

There may be two related worries with this proposal. First, the proposal looks very similar to a description theory of names along Russellian lines, which is problematic. Second, it may seem implausible that it is essential to a planet that it explains something, for could it not be that Vulcan existed, but did not explain the changes in Mercury's perihel? I will answer both worries in turn.

Since the current approach does not say anything about which theory of reference gives rise to reference realizing properties, the

---

[161] The first would be quite literally meaningless, the latter would have discernible representational properties.

present solution does not lead to the acceptance of a description theory of names. So many problems with this theory are not relevant to the present discussion. Nonetheless, there may be a worry about rigidity, for if the reference realizing properties do not include the identity-property of being identical to Vulcan, then rigidity is not guaranteed, and so it may be that the representation 'Vulcan' does not refer to the same object in every counterfactual situation. To see this, consider the following two counterfactual situations:

1) There is a planet, Vulcan, between the sun and Mercury, which explains the changes in its perihel.

2) There is a planet between the sun and Mercury, which explains the changes in its perihel, which is not Vulcan, but another planet.

The counterfactual situation in 1) is just what we would expect to be a possibility on the present suggestion, and it is, for according to the description given, something has the reference realizing properties of 'Vulcan', that is, something is the planet between the sun and Mercury, which explains the changes in its perihel. However, 2) would be impossible, provided the token representation 'Vulcan' is of the same type as the others used above, for the planet in 2) would also have to be Vulcan in virtue of having Vulcan's reference realizing properties. This potential problem ties in with the second worry, which says that it seems implausible to attribute to a planet the essential properties of being such that it explains something.

Both worries, however, may be assuaged by reminding the objector that Vulcan is not actually a planet, it is merely possible

that there is a planet by the name of Vulcan, even though it is necessary that if Vulcan exists, then it is a planet.[162]

Consider how the rigidity-worry arises in the case of existing objects: We have a definite description which in the actual world succeeds at picking out something, for example the description 'The planet closest to the sun'. We also have a name 'Mercury', which refers to the planet Mercury, the planet closest to the sun, in some suitable sense, directly. When considering counterfactual situations, the reference of these two expressions used as representations can come apart. Venus and Mercury, for example, could have switched places, and so in that counterfactual situation, the planet closest to the sun would be Venus and not Mercury. This shows that the representations ⌜the planet closest to the sun⌝, and ⌜Mercury⌝ are not representations of the same type.

In the case of Vulcan, however, there is nothing to which we could refer directly, there is nothing to which an object in a counterfactual situation could be identical to, for Vulcan doesn't exist. So no reference is determined that could come apart from the description 'the planet that explains the changes in the perihel of Mercury'. This gives us a way of understanding Barcan Marcus' remark that "identity is a relation for objects already given".[163] The notion of rigid designation makes sense only if there is an object that we can, in a suitable sense, refer to directly, for in this case the reference realizing properties of the description used to identify the object, and those of the name can come apart.

---

[162] For this to make sense, that Vulcan is not actually a planet must be read in an existentially committing way, such as not to contradict the necessity of Vulcan being a planet, if it exists.

[163] Marcus (1995) p. 213.

If there is no object we could name, there is no way for the reference realizing properties of a name and a description to come apart. Consequently there is no sense to be made of the worry that the reference realizing properties proposed above are too descriptive, for there are no two different sets of reference realizing properties, one belonging to the name and one to the description. The impossibility of there being another planet which also explains the changes in the perihel of Mercury (if it exists), but is not identical to Vulcan, is thus well justified: there are no grounds for denying that this planet is Vulcan.

So my response to Kripke-style worries about reference to mere possibilia is to admit that the reference realizing properties determined by representations of mere possibilia are not quite as determinate as those of actually existing objects. This, however, does not mean that no reference realizing properties are determined at all. Rather, some reference realizing properties need to be determined for such referential representations to figure in meaningful representations at all.

Summing up

This concludes the interlude on truth-evaluable representations involving referential representations which fail in referring to an object or objects. It prepares the ground for the discussion of further examples involving singular representations for which the puzzle about necessary existence may be thought to arise. I proposed that we should take necessarily true representation involving singular representations to not require for their truth the existence of a referent of the singular representation. Rather, we should construe the requirements for truth analogously to those given in the case of representations involving only general

representations. On this construal, the representation is a conditional claim, conditional on the existence of the object referred to, and thus true even if the object in question does not exist.

The main reasons for this position are two: First, it avoids the troublesome conclusion that everything necessarily exists. Second, it is the conditional reading on which essence attributions turn out necessarily true, while the unconditional reading does not return necessary truths.

This resolution of the puzzle of necessary existence has further advantages, because it allows us to give a unified treatment of actual objects, mere possibilia, as well as impossible objects.

2. Necessarily True Representations Involving Singular Representations

Identities

Identities involving two singular representations are commonly considered to be necessarily true. So the present theory should yield that these really turn out necessary, given reasonable assumptions about reference determination. I will start with a simple example and move on to a slightly harder one.

Simple identities involving two singular referential representations, for example the representation expressed by the sentence "Tobi is Tobi" is necessarily true. To show that the present theory yields this result, we first need the requirements for truth of the representations. I suggest we should take them to be as follows:

⌜Tobi is Tobi⌝ is true, iff the referent of ⌜Tobi1⌝ is the referent of ⌜Tobi2⌝.[164]

These requirements for truth will be guaranteed to be satisfied, and thus the representation necessarily true, just in case the reference realizing properties determined by both singular token representations are the same, that is, just in case the representations are of the same type. In this case the referent of the first will be guaranteed to be the referent of the second.

That these two referential representations are of the same type should be uncontroversial, but it can also be shown, if we assume a simple causal theory of reference: Lets assume that both singular representations are part of the same chain of uses going back to some initial baptism, where Tobi was baptized by the name 'Tobi'. If this is so, then the first representation used, ⌜Tobi1⌝, inherits its reference realizing properties from the use of the same name before and the second representation, ⌜Tobi2⌝, inherits its reference from the first. So it follows immediately that both representations determine the same reference realizing properties and that they are therefore guaranteed to have the same referent. Consequently, the requirements for truth are guaranteed to be satisfied.[165]

Clearly, however, these simple examples, while occurring once in a while, are not the only possible way in which a true identity can be expressed and also not the most informative way. More

---

[164] Here we must take care to keep the two token referential representations 'Tobi', and 'Tobi' apart, thus I add a 1 to the first occurrence in the truth-evaluable representation, and a 2 to the second.

[165] A simple causal theory along these lines is almost certainly not tenable for all names. Presently it serves well as an illustration, but as the famous Madagascar-cases show, there can be shifts of reference along a causal chain (Evans 1982).

interesting examples are sentences that, even though they merely express an identity, manage to tell us something new: The most famous example is the representation expressed by the sentence "Hesperus is Phosphorus". The first step in adjudicating whether the representation thus expressed is necessarily true according to the present theory is again to give the requirements for truth. They are analogous to the ones above:

⌜Hesperus is Phosphorus⌝ is true, iff the referent of ⌜Hesperus⌝ is the referent of ⌜Phosphorus⌝.

Once again, the representation with these requirements for truth is necessarily true, just in case the referential representations determine the same reference realizing properties, that is, if they are of the same type. The example differs from the example above, because it is not as obvious that both representations determine the same reference realizing properties.

To see how they might determine the same reference realizing properties, let us assume once more a simple causal theory of reference of the kind outlined in the semantic preliminaries. Since the representation tokens ⌜Hesperus⌝ and ⌜Phosphorus⌝ are not links in the same causal chain leading back to the same baptism, we cannot just say that the reference realizing properties are identical independently of what their content is, rather, we have to look at what the reference realizing properties are, to see whether they are indeed identical.

The reference conditions which state the reference realizing properties of ⌜Hesperus⌝ will be assumed to be analogous to those given for singular representations in the first chapter of this part. Thus I take them to be something like the following: Something is the referent of ⌜Hesperus⌝, iff it is the same planet

as the one indicated at the initial baptism of Hesperus. Equally, the reference conditions for ⌜Phosphorus⌝ are something like the following: Something is the referent of ⌜Phosphorus⌝, iff it is the same planet as the one indicated at the initial baptism of Phosphorus.

The way the identity-property is described does not tell us yet whether the same planet was present at both baptisms, as the two baptisms are distinct events. But there is a way to find out: We follow the planet from one baptism along its spatiotemporal trajectory, and see whether it is the same as the one baptized at the other baptism. Something similar arguably happened, when it was discovered that Hesperus is Phosphorus: astronomers followed the planet's spatiotemporal trajectory, and realized that Phosphorus is indeed Hesperus.[166]

Since it is in fact the same planet that was present at both baptisms, the reference realizing properties of ⌜Hesperus⌝ are identical to the reference realizing properties of ⌜Phosphorus⌝, as the identity-property which is part of the reference realizing properties is the same. And so the referent of ⌜Hesperus⌝ is the same referent as the referent of ⌜Phosphorus⌝. Consequently, the requirements for truth are guaranteed to be satisfied by how the reference of both representations is determined and the representation turns out to be necessarily true.

The desired result is conditional on the plausibility of the suggested reference realizing properties, but it does not depend on

---

[166] Of course they did not go back in time to find out and also did not have to. They knew what the referent of each name was at a certain time and place, they just didn't know that it was the same as another one observed at another time and place.

these very reference realizing properties being the reference realizing properties of the representations expressed by the words ‘Hesperus’ and ‘Phosphorus’. Rather, what needs to be the case is merely that both words are used as representations of the same type, a result that is plausible independently of the exact reference realizing properties, given that both are names for Venus.

I argued that identities, if true at all, turn out to be necessarily true on the present explanation of necessary truth, provided the referential representations involved are of the same type. However, since the truth-evaluable representations involve singular representations, the remarks from the last section apply: The representations only express necessary truths, if the existence of the referents of the referential representations is not required for the truth of the truth-evaluable representations. So a necessarily true identity does not imply the existence or necessary existence of the object which is said to be identical to itself.

Essence attributions

Essentialist truths, necessarily true representations attributing a property to an object referred to by a singular representation, have only relatively recently come back into the focus of philosophical debates, but there is a quite substantial consensus that at least some such essence attributions are necessarily true. The stock examples are attributions of kinds or origins, like “Socrates is human” or “Aristotle has the origin he actually has”, both of which supposedly express necessarily true representations.

It is controversial which attributions of essential properties are in fact true. It has, for example, been questioned whether origin is indeed an essential property of people, or whether objects are the

kind of thing they are necessarily.[167] The details of these disputes and disagreements, are not of concern to the present project. Rather, what needs to be shown is that the theory accounts for some such essence attributions, given assumptions about reference realizing properties and requirements for truth the proponent of a given essentialist thesis would accept. In other words: The theory should, at least in principle, be able to account for essence-attributions.

The stock example I will use is the representation expressed by the sentence "Socrates is human". Once more the starting point are the requirements for truth of this representation. We may suppose that they are the following:

⌜Socrates is human⌝ is true, iff the referent of ⌜Socrates⌝ is a referent of ⌜human⌝.

The truth-evaluable representation is necessarily true, if the satisfaction of these requirements is guaranteed by the reference realizing properties of ⌜Socrates⌝ and ⌜human⌝. The satisfaction of these requirements is guaranteed if the reference realizing properties of ⌜human⌝ are among the reference realizing properties of ⌜Socrates⌝, for in this case, anything that is a referent of ⌜Socrates⌝ is a referent of ⌜human⌝. So assuming for simplicity that the reference realizing properties of ⌜human⌝ are given trivially by the property of being human, being human must be among the reference realizing properties of ⌜Socrates⌝ for the representation to be necessarily true.

---

[167] For a wide ranging discussion and critique of the different essentialist theses one may hold, see Mackie (2006).

That being human is plausibly among the reference realizing properties of ⌜Socrates⌝ was already argued for in the semantic preliminaries. But let me rehearse the reasons why this is plausible: First, if being human is essential to being Socrates, then Socrates has the property of being human throughout his existence. But since reference realizing properties are also properties an object has throughout its existence, at least a necessary condition for being a reference realizing property is satisfied by the property of being human. Second, the natural kind the object is a member of is a good candidate for 'tracking' the object's spatiotemporal trajectory and as such it is a good candidate for being a reference realizing property.

So I assume that the reference conditions giving the reference realizing properties for ⌜Socrates⌝ are as follows:

Something is the referent of ⌜Socrates⌝, iff it is the same human as the one present at the initial baptism.

Once more a caveat is in order: I am not claiming that these are the actual reference conditions of the representation expressed by the word 'Socrates'. All I am claiming is that if some very simple causal theory of reference is true, and if the essentialist is right in claiming that Socrates is essentially human, then it is plausible that the reference conditions are along the lines suggested. So it is open to dispute whether the particular essentialist claim that Socrates is human is indeed true, but it follows from the assumption that a reference relation holds partly in virtue of properties of the referent that some attributions of properties to Socrates turn out necessarily true. More specifically, those essentialist claims will turn out necessarily true, which attribute to

Socrates the reference realizing properties determined by ⌜Socrates⌝.

Since it is part of the reference conditions of ⌜Socrates⌝ that he is human, and it is the trivial reference condition of ⌜human⌝ that it refers to whatever is human, whenever ⌜Socrates⌝ refers to something, so does ⌜human⌝. Consequently ⌜Socrates is human⌝ is necessarily true according to the account, provided the requirements for truth as well as the reference conditions are as specified.

As in the case of identities, the necessary truth of ⌜Socrates is human⌝ only follows on the account, if the requirements for truth do not require the existence of a referent of ⌜Socrates⌝ for the truth of the truth-evaluable representation. For it does not follow from the reference realizing properties of ⌜Socrates⌝ that there is something that has these properties. So the requirements for truth should be formulated more perspicuously as follows:

⌜Socrates is human⌝ is true, iff, if there is a referent of ⌜Socrates⌝, then this referent is a referent of ⌜human⌝.

This is, I explained earlier, a desirable feature of the account, as it avoids the necessary existence of everything that has an essential property, but still accounts for the necessary truth of essence attributions.

I argued that the present theory can account for the necessary truth of essence attributions, and that this is not markedly different from the account of the necessary truth of representations involving only general representations: In both cases the requirements for truth are guaranteed to be satisfied by

the reference realizing properties of the constituent referential representations, provided these requirements for truth do not require that the referential representations have a referent.
What is left to account for in this chapter are mathematical truths, which are commonly also taken to be necessarily true. This is the task of the next section.

Mathematical truths

Mathematical truths present a somewhat special challenge to the present theory, as it is among other things controversial what the referents of number-terms are, whether numbers exist at all, and how the meaning of the mathematical operators should be spelled out. I cannot and will not solve all these problems here, I will merely propose one way in which a suitable subset of mathematical truths may turn out to be necessarily true on the present account. This should give the reader some idea of how the account can be used to cover further mathematical truths. The basic assumption concerning mathematic truths will be a form of realism about them: mathematical truth is a regular kind of truth. Mathematical representations are true in virtue of the referents of the referential representations standing in the appropriate relationships to each other. Thus mathematical truth is not to be equated with, for example, provability, etc. A consequence of this account will be that mathematical entities like numbers do not exist necessarily, even though mathematical truths, even those apparently involving existence claims, may turn out necessarily true. I will, after outlining the basic account, defend this possibly controversial consequence.
Paradigmatic mathematical statements are identities, for example the following: 2=2, 2+3=5, 21/7=3. But not all mathematical

statements are of this sort. There are different relations that hold between numbers, for example the greater than or smaller than relation, giving rise to statements like the following: $2<5$, $4>3$. Apart from these, there are statements apparently stating the existence or non-existence of some number: There is no greatest natural number; there are two prime numbers smaller than 5; there are two numbers which solve the equation $x^2+5x+5=0$, and so on. All of these statements seem to express representations that are necessarily true. To show how the current theory may yield the desired result, that is, the necessary truth of mathematical representations, I will move from the simple examples of identities to the more complicated quantified statements, all along assuming that the number terms do represent numbers, whatever numbers may be.[168]

The simplest example are identities involving just one number term on each side of the identity sign: "2=2". These examples can be treated in the same way as regular object-identities: Whatever determines the reference of the representation ⌜2⌝, the reference determining properties of the representation on each side of the identity determine the same reference realizing properties and so both are guaranteed to refer to the same thing: the number 2. Given the requirements for truth for identities outlined above, ⌜2=2⌝ turns out to be necessarily true on the present account.

This example is easy, but what about more complicated mathematical representations? For example the representation

[168] I will not have much to say on what the referents of number-words are. One may, for example, think of numbers as properties of collections, or properties (cardinalities) of sets of objects, and of our number words as referring to these properties. A possible suggestion, I am sympathetic to, is made in Moltmann (2013)

represented by “2+2=4”. The basic form is still an identity, but this time the left hand side apparently involves two singular representations as well as ‘+’. Since what is supposed to be true is an identity, and identities are true if the representations on either side of the identity have the same referent, we should take the entire expression ‘2+2’ to express a representation referring to the number four. We can think of 2+2 as a rigid description of the number 4: Just like we can pick out people by invoking a position relative to other people, “The person next to John”, we can pick out numbers by invoking their position in the number series relative to others. This much should be uncontroversial, but how does this give rise to the necessity of the representation? For the representation to be necessarily true, it needs to be the case that ⌜4⌝ and ⌜2+2⌝ have the same reference realizing properties, that is, that they are representations of the same type.

We may, roughly following Peano, think of the reference conditions of natural numbers as being given relative to the number 0 in terms of being the successor of 0. Following this thought, the reference conditions of ⌜4⌝ are that the referent of ⌜4⌝ is the fourth successor of 0. Similarly, we may think of the reference conditions of ⌜2+2⌝ as being given in the following way: the function x+2 picks out the second successor to the number which stands for x and ‘2’ picks out the second successor of 0. So the whole term picks out the second successor of the second successor of 0. Consequently the reference conditions for the representation ⌜2+2⌝ are that the referent is the fourth successor of 0. It is now obvious how the identity is bound to be true: the reference realizing properties of both, ⌜4⌝ and ⌜2+2⌝ are

identical, and so the requirements for truth are guaranteed to be satisfied.

So if the reference conditions given here are plausible in outline, at least basic identities in mathematics can be treated analogously to other identity statements, and turn out necessarily true on the current theory.

One may be slightly concerned about a feature of this analogy, however. In the case of identities of ordinary objects, the necessity which the account could and should account for was necessity conditional on the existence of the object that is supposed to have a property necessarily. In the mathematical case, however, this may be taken to be too weak, for it is often taken for granted that numbers and other abstract mathematical objects, exist necessarily.[169]

The necessary existence of numbers, however, is not a consequence of the present account, and could not be, for no reference condition can guarantee its own satisfaction. This may not cause much trouble in the case of identities, for at least they do come out as necessarily true, if conditional on existence, but other mathematical representations seem to be existence claims, and so apparently require the necessary existence of numbers for their necessary truth.

I will tackle this problem head on, and argue that it is an acceptable consequence of the current theory that mathematical objects do not come out as existing necessarily. What does turn out to be necessary, however, is, as in the case of essence

---

[169] For arguments in favor of the necessary existence of numbers see Tennant (1997), as well as Hale and Wright (1994). For an opposing view see Field (1989) and Hellman (1989). Field as well as Hellman, however, have much stronger views than me, for their view is that numbers do not even actually exist. The present theory is committed to no such thing.

attributions, a conditionalized understanding of mathematical representations. This may be seen as a partial validation of mathematical if-then-ism.[170] It should be noted, however, that currently no stance is taken on whether numbers actually exist, or on how *all* mathematical statements should be understood. As in the case of ordinary objects, some mathematical statements may also be interpreted as requiring the existence of numbers for their truth, but in so far as they do, they are not necessarily true, even if they may be actually true in virtue of the actual existence of numbers. So the claims here should be read as saying that if mathematical representations are to be necessarily true, then existence claims of particular numbers are to be understood as conditional on the existence of numbers.

Let us look at a concrete example of a true mathematical existence claim: There is a natural number which is prime and smaller than 3. The most natural understanding of this statement is that it states the existence of a natural number which satisfies two conditions, being smaller than 3, and being a prime number. So the requirements for truth, if taken at face value, should be something like this:

⌜There is a natural number which is prime and smaller than 3⌝ is true, iff there is a referent of ⌜natural number⌝ that is both a referent of ⌜prime number⌝, and ⌜smaller than 3⌝.

---

[170] Putnam (1967) coined the term. The view is quite controversial as it seeks to give a translation-manual of mathematical statements into statements conditional on the existence of numbers. The present view endorsed here need not make such a strong and controversial claim. The present commitment is merely that in so far as mathematical representations are necessarily true, they must not require the existence of any object for their truth. It is left open that some mathematical statements, especially in applied mathematics, do require the existence of numbers for their truth (whatever numbers may be).

However, as noted, these requirements for truth involve an explicit existence claim and thus cannot follow from the reference conditions of the constituent representations. So there is no way for these requirements for truth being guaranteed to be satisfied by the reference realizing properties of the referents.

Consequently it may seem that some mathematical truths do not turn out necessarily true on the present theory. This verdict, however, is only partially correct, for it is indeed not the case that the necessary existence of mathematical objects, or any object for that matter, follows from the present theory. However, a mathematical truth closely related to the existence claim does turn out to be necessarily true on the present theory: The claim that if the natural numbers exist, then there is a natural number among them that is both prime and smaller than three. Since the antecedent of this claim can be taken for granted when working on mathematical questions, the non-necessary unconditional existence claim can be viewed as a kind of short hand for the conditional claim which is necessarily true.

To see how the conditional claim turns out necessarily true, it helps to look back at essence attributions to ordinary objects. The representation expressed by the sentence “Socrates is human”, requires for its truth that if there is a referent of ⌜Socrates⌝, then it is a referent of ⌜human⌝. This, it was claimed above, is guaranteed to be the case in virtue of the reference realizing properties of ⌜Socrates⌝ and ⌜human⌝. Since the requirements for truth are conditional on the existence of Socrates, it does not follow from them that necessarily, Socrates exists and is human. However, a slightly different existence claim follows from the essence-claim: the claim that if the humans exist, then there is

something which is Socrates and which is human.[171] This claim follows, because it simply strengthens the antecedent of the conditional which is necessarily true.

In the mathematical case we can view the conditional that all numbers exist as a presupposition, or common assumption, of the mathematical enterprise, which is not made explicit. So we can view the claim that there is a prime number smaller than 3 as making the conditional claim, that if the natural numbers exist, then there is a prime number smaller than 3. To show that this really is necessarily true, we can proceed analogously to the Socrates case: We show that a conditional claim with a weaker antecedent is true and strengthen the antecedent. The claim which can be shown to be necessarily true is that if the number two exists, then it is a natural number that is both smaller than three, and prime, from which it follows that if the natural numbers exist, then there is a number that is both smaller than three and prime. First, we need the requirements for truth for this representation:

⌜The number two is a natural number which is prime and smaller than 3⌝ is true, iff if there is a referent of ⌜2⌝, then it is a referent of ⌜natural number⌝, ⌜smaller than 3⌝, and ⌜prime number⌝.

What has to be shown next, is that these requirements for truth are indeed guaranteed to be satisfied by the reference realizing properties.

---

[171] It is slightly unclear how this particular existential claim should be understood in the present framework. Above existence claims were treated in the requirements for truth as requiring that there is a referent for a representation. Here the existence claim is much more general, for it says not merely that some human exists, but that all humans exist. This is analogous to the claim that the natural numbers exist, meaning that all natural numbers exist. I suggest we treat this as requiring that there is a referent for every possible type of representation of a natural number, or of a human being respectively.

The reference conditions stating the reference realizing properties for ⌜2⌝, are, we may assume in analogy to ones suggested above, that the referent of ⌜2⌝ is a natural number that is the second successor of 0. Further, we need reference conditions of ⌜smaller than 3⌝, and ⌜prime number⌝. Here are the suggestions following the standard definitions:

Something is a referent of ⌜prime number⌝, iff it is a natural number greater than 1, and divisible only by itself, and 1.

Something is a referent of ⌜smaller than 3⌝, iff it is a natural number not equal to 0, such that if added to the referent, it is equal to 3.

Now we can see how the requirements for truth are guaranteed to be satisfied: ⌜smaller than 3⌝, provided that natural numbers exist, refers to 0,1, and 2. And ⌜prime number⌝ has the number 2 as a referent, if it exists. So, if something is the referent of ⌜2⌝, it is the referent of ⌜prime number⌝,⌜smaller than 3⌝, and (trivially) of ⌜natural number⌝.

Since the truth of the representation ⌜The number two is a natural number which is prime and smaller than 3⌝ is sufficient for the truth of the claim that if the natural numbers exist, then there is a natural number which is both smaller than 3 and prime, it also turns out necessarily true.

These examples are not very complex, and merely a first suggestion of how the present theory may be able to account for the necessity of mathematical truths. The examples at least show that the present theory has the resources to account for the

necessary truth of mathematical representations conditional on the existence of numbers. Before moving on, I discuss one final worry:

One may be concerned that in the last example above, I am endorsing a form of mathematical if-thenism, or fictionalism, the view that we should understand all mathematical statements as conditional on the existence of numbers. But if-thenism is not a particularly popular thesis, for it seems to go against the most natural reading of ordinary mathematical statements. We just do not seem to conditionalize in the way advocated when making such statements.[172]

It should be noted, however, that I do not make such a strong claim. It is consistent with the discussion above that ordinarily mathematical statements should not always be understood conditionally. Even if understood unconditionally, the statements may be true, because numbers do actually exist. However, these statements will not be necessarily true, for the current theory of necessity denies the necessary existence of anything. What is necessarily true, is the conditional claim and combined with the truth of the antecedent, the truth that natural numbers exist, the existence involving claim follows.

But why, one may ask, should we accept that the requirements for truth of mathematical truths should change from requiring the existence of numbers to being conditional on the existence of numbers, if we put a modal operator in front of it? After all, we do

[172] This straightforward criticism seems to be considered by some to be uncontroversial enough to dismiss the position without further argument. See for example Balaguer (2015). But see Moltmann (2013) for a detailed analysis of our everyday number talk.

use the same form of words in the modal and in the non-modal context.

The answer is that a 'weak' reading of the necessity, conditional on existence, is the most plausible one for two reasons. First, we may draw an analogy to other cases of weak necessities, for example identity claims of and essence attributions to ordinary objects. Here the conditional reading, even though not explicit in the form of the statement, is quite natural, for we want to preserve the thought that ordinary objects exist only contingently, a truth inconsistent with the necessity of the unconditional reading. This gives us reason to think that the necessity of identity claims are to be understood as not requiring the existence of referents of the singular representations contained in them for their truth. If this applies to ordinary objects, it seems that it should not be excluded in the mathematical case. To the contrary, it seems that unified requirements for the truth of identity claims are quite desirable.

Second, it is only conditional requirements for truth that can be guaranteed to hold in virtue of the reference realizing properties. So the only mathematical claims that do turn out necessarily true are the ones not requiring the existence of numbers. This is a consequence of the present account of necessity which implies that nothing exists necessarily, for no reference condition can guarantee its own satisfaction.[173] So given the current theory is true, a necessity claim involving a mathematical truth should be read as a weak necessity, conditional on the existence of numbers, if we want to be able to interpret it as true.

---

[173] This is not to say that existence cannot be an essential property of some object, it is only to say that this does not imply the existence of the object which has existence as an essential property.

This second argument depends on the plausibility of the proposed theory, and so my opponent may well try to use the impossibility of accounting for necessary existence as an argument against the theory on the grounds that mathematical objects do exist necessarily. This move, however, will only have sufficient support, if a theory can be presented on which necessary existence can be explained and at the same time a plausible resolution for the problem of necessary existence can be found.

To conclude, the current theory implies that mathematical truths, in so far as they require for their truth the existence of numbers, are not necessarily true. However, if read conditionally, as not requiring the existence of numbers for their truth, mathematical truths do turn out to be necessarily true according to the present explanation. This weak reading is plausible by analogy to identity claims involving ordinary objects, and does not commit the proponent of the current theory to an overly general view about how all mathematical claims should be understood.

Summing up

The above examples have illustrated how the suggested explanation of necessity applies to a range of necessary truths, and how the framework allows for a unified treatment of a number of potential puzzles surrounding necessary existence, possibilia, and impossible objects. I will here consider how far the discussion has proceeded, how much this should contribute to our confidence in the proposed explanation of necessity, and what remains to be shown in the following chapters.

I developed a general framework in which the core concepts can be understood. This framework provided the basic concepts for an explanation of necessary truth: Reference realizing properties, and

requirements for truth. The reference realizing properties are those properties of a referent which are constitutive of the reference relation between a referent and a referential representation. What these properties are, or would have to be for a reference relation to hold, is determined by the reference determining properties of a referential representation. The requirements for truth are the requirements a truth-evaluable representation places on the referents of the constituent referential representations for its truth.

In some cases, the reference realizing properties determined by the referential representations are such that they guarantee the satisfaction of the requirements for truth of the truth-evaluable representation. In such a case the truth-evaluable representation cannot fail to be true, it is necessarily true. So the appropriate relationship between reference realizing properties and requirements for truth explains necessary truth in cases where the appropriate relationship holds.

The examples in this chapter were supposed to illustrate how the explanation of necessity works in a range of cases, but they were also supposed to raise confidence in the extensional adequacy of the theory, that the proposed explanation covers the standard cases of the most fundamental kind of necessity: metaphysical necessity.

Along the way I argued that the general outlook provided by the theory allows for a unified treatment of existing objects, merely possible 'objects', and impossible 'objects'. A puzzle about necessary existence, arising from the existence of necessary truths about contingently existing objects is quite naturally resolved along the standard lines, because only necessary truths about an

object that do not require the object's existence for their truth turn out necessarily true on the present theory of necessity.

I take it that the discussion of the examples has shown that there is at least no principled obstacle to the proposed explanation covering all kinds of representations that have commonly been taken to be metaphysically necessary. As such it is not limited to simple subject-predicate constructions like the Kantian approach, and it is also not limited to necessities involving merely general referential representations, but can, at least in principle, account for essence-attributions and identities. The latter is made possible by the fact that in the general framework, there is no difference in kind between the reference relation that holds between a general representation and its referents and a singular representation and its referent, as in both cases the reference relation holds constitutively in virtue of properties of the referent and properties of the representations.

To demonstrate how the present theory accounts for these diverse necessary truths, I made some assumptions about the particular reference realizing properties. These assumptions were made along plausible lines, utilizing some common assumptions about reference-determination. However, they did remain assumptions, and I would be hard pressed to pretend that these assumptions are fully correct for the representations used. So it may be that some of the assumptions do turn out false on closer inspection of what really determines reference for these representations.

This, however, should not be damaging to the overall account, for while it may show that some particular representations do not turn out necessarily true, a true account of how reference is really determined will yield a recipe for constructing a host of necessary truths. So what a failure of these assumptions should prompt, is a

revision of our preconceptions about what is necessarily true, rather than a rejection of the proposed explanation.

The reader may not yet be quite convinced by this assessment, so the next chapters are designed to bolster trust in the theory by showing that it is best conceived of as an account of what has been known as metaphysical necessity.

In the next chapter, I will argue that the necessity in question can and should be taken to be the basic kind of necessity from which other, weaker alethic necessities such as nomic necessity can be understood. Following this discussion of how the explanation fits with the varieties of necessity, I will compare the present explanation to essentialism, which is uncontroversially a theory of metaphysical necessity. I argue that the present proposal can be viewed as a form of essentialism, one which does not take essences to be primitive, but gives an account of what essences are: the essential properties of an object are the reference realizing properties determined by the singular representation referring to it.

Chapter 7: Varieties of Necessity

In the last chapters I proposed an explanation of what has been called metaphysical necessity by invoking an appropriate relationship between reference realizing properties and requirements for truth of a representation. Metaphysical necessity, however, is not the only kind of necessity that needs explaining, and so in the following I will show how the present theory may explain other varieties of necessity such as nomic or practical necessity.[174] I will argue that a classical relativization strategy is most naturally combined with the overall account to explain necessities weaker than metaphysical necessities, and that there are good reasons for adopting it. Stronger necessities than metaphysical necessity, for example logical necessity, can be accounted for by restriction. To discern the challenges an account of other kinds of necessity faces, I will first have a look at the structure of alethic modality. Second, I demonstrate how the current theory accounts for the different kinds of modality before two challenges are confronted. The first challenge is to give a principled reason for taking metaphysical necessity (in the sense in which the necessity in question explained by the current account is metaphysical necessity) as the fundamental kind of necessity from which the others are explained. The second is a challenge, for example voiced by Kit Fine and Marc Lange, who claim independently of each other that natural necessity is not metaphysical necessity relative to the laws of nature, but a

[174] The only kinds of necessity which will be considered are well-behaved alethic necessities, those for which it holds that if necessarily p, then p, which are weaker than metaphysical necessity. This excludes all non-alethic necessities, but also some alethic necessities such as epistemic necessity, which is, at least partially, stronger than metaphysical necessity.

different kind of necessity, on the grounds that a relativization strategy does not capture the distinctive necessity of natural or nomic necessity.[175]

1. The Structure of Modality

Different kinds of alethic necessity can be ordered according to their strength. Metaphysical necessity is a fairly strong kind of necessity, while practical or nomic necessity are weaker in taking more propositions to be necessarily true.[176]

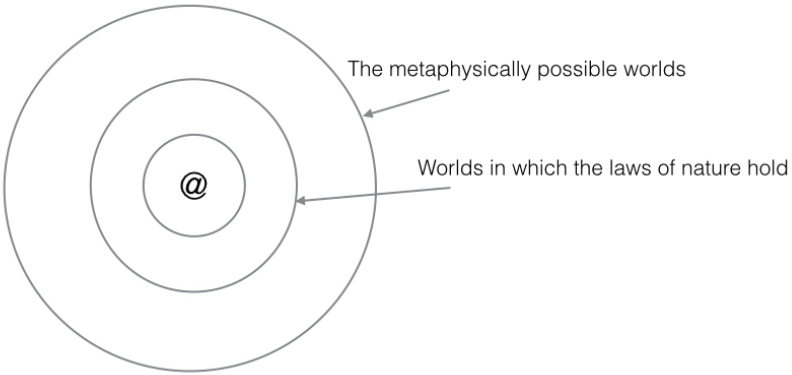
The structure of modality can be nicely represented in terms of possible worlds: Imagine that all possible worlds there are, represent all the possibilities there are in the widest sense. Whatever is true in all of them is, let us assume for the moment, metaphysically necessary, and whatever is true in some of them is metaphysically possible. When talking about these possible worlds and quantifying over them, we can restrict our quantifiers and only talk about a subset of all these possible worlds. In this subset, there are also some truths which are true in all of the worlds of the subset, and some that are true only in some of them.

---

[175] Fine (2002), Lange (2009).

[176] Talk of weaker or stronger kinds of necessity as well as talk of restricted and unrestricted possibilities can sometimes be slightly confusing, for the terms are used in different ways (In 2D-semantics, for example, the distinction means something quite different). So here is how I will use the terms: Some kind of necessity N1 is strictly stronger than another N2, iff all truths that are necessary according to N1, are a proper subset of all truths which are necessary according to N2. Corresponding to this definition of the strength of a kind of necessity is a distinction between more or less restricted kinds of possibilities. A possibility of kind P1 is more restricted than a possibility of kind P2, iff all truths which are possible according to P1, are a proper subset of all truths which are possible according to P2. Given the interdefinability of necessity and possibility relative to the kind of necessity at issue, restriction is related to strength in the following way: The stronger the necessity the more unrestricted the corresponding possibility.

These are the truths that are necessary or possible respectively relative to that subset. We can find many such restrictions on the set of all possible worlds, but the ones we might care about most are the ones that are made according to certain principles. So for example, if we care about what is possible or necessary given the laws of nature, we choose that subset of all possible worlds, in which our laws of nature hold, and then go on to evaluate for necessity and possibility on the subset restricted in this way. This same procedure, just with different subsets of all the worlds picked out, works for other alethic modalities as well, for example for practical necessity, and gives us a nice picture in which we can order the necessities we care about according to their strength:[177]



[177] Note that it may be that there is not quite such a neat ordering, as the sets of worlds to which we restrict our quantifiers may not all be subsets of each other. But this complication need not concern us here. What is important is merely that the sets of worlds we restrict ourselves to, is a subset of all metaphysically possible worlds.

The picture may suggest that all kinds of weaker than metaphysical necessities are neatly contained in each other, but this need not be the case. Even though it is plausible that the nomic necessities are a subset of the practical necessities, for example, and that metaphysical necessities are a subset of both, it may be that other weaker kinds of necessity do not exhibit this neat containment structure, depending on what they are relativized to.

Given this quite general structure of necessity, with metaphysical necessity at its core, the question arises how we can account for the other kinds of necessity. The picture already offers suggestions of how this may be done: If we have an explanation for one kind of necessity, we may define the others either by restriction or relativization.[178] One may, for example, take the metaphysical necessities as the core necessities and define natural necessity and possibility by relativizing to the laws of nature: Natural possibilities are those truths which do not yield metaphysical impossibilities if conjoined with the laws of nature. One may, however, also take some weaker kind of necessity as primitive, for example natural necessity, and define the stronger necessities by restriction. In this spirit, we may conceive of the metaphysically necessary truths as the subset of the naturally necessary truths which are, in some sense, metaphysical truths. The two strategies may of course also be combined in accounting for different kinds of necessity.

Both strategies can be used to defend modal monism, the view that all (well behaved) alethic modal truths have, in a sense to be

[178] For a further exploration of the different strategies, see Fine (2002).

specified, the same source.[179] For now, modal monism will be assumed, because it is the dominant view, and alethic modalities do seem to have a lot in common. Challenges to this assumption will be discussed in due course.

2. Explaining the Varieties of Necessity

The fundamental kind of necessity is, I will argue, metaphysical necessity as explained by the current account. Weaker kinds of necessity, such as nomic necessity, can be explained by relativization, and stronger kinds can be explained by restriction.

Let us start with weaker kinds of necessity and take nomic necessity as an example. Nomic necessities are those truths, which are metaphysically necessary, given that the laws of nature hold. So we may define them by relativizing to the laws of nature:

> *(Def. nomic necessity -)* A representation R is nomically necessarily true, iff the conjunction of the negation of R with the laws of nature is a metaphysical impossibility.

or to put it in positive terms:

---

[179] The term 'modal monism' is used in different ways in related debates. In particular in the debate on the epistemology of modality, modal monism is the position that the space of epistemically possible worlds does not exceed the space of metaphysically possible worlds. The difference between epistemic and metaphysical necessity, on this view, comes down to different ways in which we may consider other possible worlds. We can consider them either as actual, or as counterfactual. Modal monism in this sense is commonly associated with modal rationalism, and defended in some detail by David Chalmers (2012), who uses a two dimensional semantic framework to spell out his ideas.
Modal monism in the present metaphysical discussion is not to be confused with this use in the epistemological debate. The kind of monism at issue presently concerns the ground of necessity. Accordingly, I leave the treatment of epistemic necessity to the side and merely offer some remarks on how to deal with it at the end of the present chapter.

> *(Def. nomic necessity +)* A representation R is nomically necessarily true, iff the conditional containing R as a consequent and the laws of nature as antecedent, is metaphysically necessary.

Let me give an example to see how this works. Let's assume for simplicity that there is only one law of nature and that it is the following: All swans a white. Then it should turn out to be nomically necessarily true, that if Johanna is a swan, then she is white. By our definition of nomic necessity, this is true, if and only if the following conditional is metaphysically necessary: If all swans are white, then, if Johanna is a swan, then she is white. Now it is fairly obvious that this conditional is metaphysically necessarily true, for the consequent is just an instance of the antecedent, and thus it is guaranteed that whenever the antecedent is true, the consequent is.[180]

This way of explaining nomic necessity can be straightforwardly generalized to other weaker kinds of necessity by varying the antecedent of the metaphysically necessary conditional. So we may use the following schema to arrive at different kinds of weak necessity:

> *(Weak Necessity Schema)* A representation R is X-necessarily true, iff the conditional containing R as a consequent and the conjunction of all X-truths as antecedent is metaphysically necessary.

---

[180] I will not go into the details here. It has been shown above how the necessity of truths of logic can be explained on the present account. The explanation can be applied in the present case as well.

The schema can be used quite generally to define weaker kinds of necessity, as there is no special restriction on the X-truths that enter into the antecedent, except that they need to be true.

While weaker kinds of necessity than metaphysical necessity are often of interest, they are not the only other kinds of necessity there are, for there may be necessities that are stronger than metaphysical necessity. One example is logical necessity. Logical necessities are metaphysically necessary, but not all metaphysically necessary truths are logically necessary. So we may define the logically necessary truths by restriction:

Logically necessary truths are that subset of the metaphysically necessary truths, that are guaranteed to satisfy their requirements for truth in virtue of only the reference realizing properties of the logical relations.[181]

Other more or less principled restrictions may be found as well to pick out a limited number from all metaphysically necessary truths. So in principle the present theory can account for the varieties of necessity in the standard way, by proceeding form metaphysical necessity by relativization or restriction.

3. The first Challenge: Finding the Right Starting Point

Given that the structure of necessity is as outlined above and that the present account can be used to explain weaker as well as stronger necessities in terms of metaphysical necessity, it remains to be shown that the right starting point for this explanation is

[181] This proposal is incomplete, as the treatment of logical truths in the example section showed that we should only conceive of statements of predicate logic as stating logical relations between propositions, that is, relations which hold merely in virtue of the truth values of the propositions. Either way, however, there is surely a way to pick out the relevant representations, which are to be considered for the appropriate restriction.

metaphysical necessity. To defend this view, it will have to be shown both that there is no stronger kind of necessity which can plausibly be taken as the starting point of the relativization and that there is no weaker necessity which should be taken as the starting point for restrictions.

The theory-independent considerations about the structure of necessity do give us some reason to take metaphysical necessity to be the right starting point for an account of weaker kinds of necessity. I will first give an argument to the conclusion that no necessity stronger than metaphysical necessity, as explained by the present theory, is a good starting point for an account of the varieties of necessity, and then proceed to argue that no necessity weaker than metaphysical necessity should be the starting point for a theory seeking to explain the varieties of necessity by relativization or restriction.

An example for a stronger kind of necessity than metaphysical necessity is logical necessity. It is stronger, because there are some truths that are metaphysically necessary, which are not logically necessary. Examples are “All bachelors are unmarried” and “If Socrates exists, he is human”. Both of these truths are, we may assume, metaphysically necessary, but they are commonly not taken to be logically necessary, because we cannot establish their truth merely by considering their logical form. It is quite clear that if logical necessity is alethic, then every logical necessity is metaphysically necessary, but not the other way

around.[182] Since necessity and possibility are duals, we can define "possibly p" as "not necessarily not p". So what is not necessarily not logically necessary, is logically possible. Consequently, it is logically possible that if Socrates exists, then he is not human and it is logically possible that bachelors are married.

However, there is absolutely no way in which it can be true that if Socrates exists, he is not human, and there is no way in which it can be true that bachelors are unmarried, even if this is not excluded just by the logical form of the sentence. So (narrow) logical possibility does not in any way guarantee that what is logically possibly true, is possibly true in an intelligible sense.

What does this show? It shows that logical necessity is not the fundamental kind of necessity, because to arrive at what is possibly true in an intelligible sense, we need to relativize logical necessity to a number of metaphysical or conceptual 'laws'. This would be undesirable, as we should expect an account of the fundamental kind of necessity to also tell us what is, in some intelligible sense, possibly true.

This argument is not a knockdown argument, as there may be disagreement among modal monists as to what should be required of the fundamental kind of necessity. But it seems like a reasonable requirement that the notion of possibility it yields, should give us a sense of possible truth that is intelligible. If

---

[182] It may be doubted that logical necessity is alethic, because some apparently formally valid arguments are not truth preserving. Consider 1) Jakob fought god at Peniel, 2) I am Jakob, so 3) I fought god at Peniel. 3) is obviously not true, even if we take the premises to be true, but just considering logical form, it is valid. Of course, in artificial languages it is typically not allowed that the same word can be used as a name for more than one object, and so the argument cannot get off the ground, but this is already a restriction on the meaning of the names used, and so form alone cannot guarantee truth-preservation. This complication, however, will not be considered further.

effective, the argument does not just speak against taking logical necessity to be the fundamental kind of necessity, but against taking any kind of necessity stronger than metaphysical necessity to be fundamental, as there will be some truths to be counted as possible in that sense which absolutely cannot be true.[183] This suggests that we should not take any notion of necessity stronger than metaphysical necessity, in the sense explained by the current theory, to be the fundamental notion of necessity.

The reader may object to this argument on the grounds that it begs the question against anyone seeking to take a stronger notion of necessity as his starting point. For according to the philosopher taking, for example, narrow logical necessity to be fundamental, it will turn out to be genuinely possible that, say, Socrates is not human, and that bachelors are married.

This would be true, if we would consider this question independently of the available explanation of metaphysical necessity offered and in particular independently of the framework offered for this explanation. For given this framework, there is no way for a representation representing Socrates to not be human to be true and there is no way for a representation to be true that represents bachelors as married. So no sense can be made of the possible *truth* of such a representation and consequently the argument that the kind of possible truth corresponding to logical necessity remains unintelligible stands. The proponent of a stronger kind of necessity as fundamental would have to show how the correlative notion of possibility is still a notion of possible *truth.* I doubt that this can be done.

---

[183] Another prominent kind of necessity stronger than metaphysical necessity is for example broadly logical necessity, which includes, in addition to the narrowly logical truths, the conceptual truths.

What remains to be shown is that no notion weaker than metaphysical necessity should be taken as fundamental. If we take a weaker notion of necessity than metaphysical necessity to be fundamental, there will be some truths which are (intelligibly) possibly true, but impossible in the weaker sense of possibility. This is the flip-side of the objection against taking stronger than metaphysical necessities to be fundamental, and it suggests that taking metaphysical necessity to be fundamental yields a more uniform account of all weaker kinds of necessity. While this again does not constitute a knockdown argument, it is suggestive.

Let us assume that natural necessity is weaker than metaphysical necessity. If so, then there are some natural impossibilities which are still possible in an intelligible sense. It is, for example, not naturally possible to travel faster than the speed of light, but it still seems to be possible, in the metaphysical sense, to travel faster than the speed of light. If natural necessity is the most fundamental kind of necessity, then it would seem to be unclear in which sense it is possibly true to travel faster than the speed of light. The sense in which it would be possibly true, would presumably be that it is not incompatible with a subset of the naturally necessary truths. But, even though it is true that the metaphysically possible truths are not incompatible with a subset of the naturally necessary truths, this does not seem to capture the sense in which it is possible to travel faster than the speed of light, for this seems to be a genuinely possible truth, not just compatible with a subset of all fundamentally necessary truths, but incompatible with all of them taken together.

The general lessons to be learned from these remarks is the following: If one wants to be a monist about necessity, then one should take as one's starting point the kind of necessity which is

weak enough to have a correlative notion of possibility that only includes truths that can possibly be true in an intelligible sense, and it should be strong enough to capture the sense in which it is possible in the widest sensible sense for something to be the case. The kind of necessity which satisfies these two conditions, I have suggested, is metaphysical necessity as it is captured by the current account of metaphysical necessity. All weaker kinds of necessity can then be captured by relativization, and all stronger kinds by restriction in the way outlined above. This will yield an account of the varieties of necessity that is elegant, simple, and captures the sense in which metaphysical necessity is fundamental to all necessities.

One may at this point object to the latter part of the argument on the grounds that it is misguided to think that it is, in an intelligible sense, possible to travel faster than the speed of light, as we should take the laws of nature to be metaphysically necessary, and therefore take it to be in no sense possible to travel faster than the speed of light.[184] On this view, the class of metaphysical possibilities would be much narrower than suggested above. The first thing to note in response to this worry is that even if it was true that the natural necessities are all metaphysically necessary, the observations above would hold, for in this case metaphysical necessity would be coextensive with natural necessity, and so natural necessity would be the strongest kind of necessity whose correlative notion of possibility represents genuine possibilities and the weakest kind such that the correlative notion of possibility includes all possibilities which are in the most inclusive, intelligible sense possible. The only disagreement with the

[184] The view that the laws of nature are necessarily true is for example defended by Shoemaker (1998)

proponent of the metaphysical necessity of natural laws would be about which truths to count as metaphysical necessities. But this is a dispute that does not need to be decided here, as these substantive question will have to be decided by a more in depth inquiry into the referential representations used to refer to the objects and properties appearing in the laws of nature. It is perfectly compatible with the present account that the laws of nature, or at least some of them, do indeed turn out to be metaphysically necessarily true. For this to be the case, it would have to be part of the reference conditions for, say, having a mass m, that the object which has mass m behaves according to the inverse square law. It is possible that we do have representations which have such demanding reference conditions, but it is, I submit, not very likely to be true that our actual use of terms of the type 'having mass m' refer only to objects that obey the inverse square law.

This is a side-issue to the current debate, however. The purpose of this section was to defend metaphysical necessity, the kind of necessity explained by the present theory, as the monist's best starting point for accounting for the varieties of necessity. This is in principle compatible with the natural necessities being coextensive with the metaphysical necessities.

The best starting point for the monist about the varieties of necessity is, I hope to have shown, metaphysical necessity. From this principled starting point, both stronger and weaker kinds can be defined by restriction and relativization respectively. The assumption throughout was that monism is true. This, however, has been contested by eminent voices, who have argued that different kinds of necessity have different sources and should

therefore be kept distinct.[185] This challenge is discussed in the following.

4. The Second Challenge: Different Sources of Necessity?

Monism about the varieties of necessity was and is the standard view. Possible world theories, for example, start with the class of all metaphysically possible worlds and explain weaker necessities with contextually induced variations in the worlds we quantify over. This is not an accident. The monist view is popular, because it nicely captures the structure of alethic modality as it was outlined above: It seems to be the case that the metaphysical necessities are a subset of the nomic necessities, which in turn seem to be a subset of still weaker necessities such as practical necessity. And correlatively, the naturally possible worlds are a subset of the metaphysically possible worlds. For the monist this structure is readily explained, for it is implied by the account: nomic necessity, for example, is arrived at by relativizing to a smaller portion of all possible worlds according to some criterion. The current theory, even though it does not invoke possible worlds, equally easily explains this structure, since it also arrives at weaker necessities by relativization.

If one were to deny monism, one would either have to deny this structure or alternatively find a way to explain why it seems to be the case that all metaphysical necessities are also nomic necessities. It is at least not obvious how such a strategy may proceed, if the source of necessity is different for different kinds of necessity, so there is reason to be a monist by default.

---

[185] Most notably Fine (2002). For detailed critical discussion of Fine see Leech (2015).

Still, monism has been challenged, and convincingly so. Kit Fine offers two somewhat related challenges to it. First, he tries to show that the structure of modality is not as would be predicted by an account which proceeds by relativization of metaphysical necessity and thus that relativization-accounts are extensionally inadequate.[186] Second, he suggests that no relativization-account will capture the special nomic force of natural laws. The latter of these criticisms has been made by a number of other authors as well, notably by Marc Lange.[187] As the challenges are connected, I will discuss them together in what follows.[188]

Fine's counterexamples proceed by attempting to show that worlds in which laws hold that do not have a subject matter to be governed by these laws, lead to contradictory assessments of what is possible in these worlds.[189] I simplify the examples somewhat, but the central point, as well as how to challenge it, will come out quite clearly: Consider two worlds, one (w1) governed by, say, the inverse square law, a second (w2) governed by the inverse cube law. It seems nomically possible for both worlds to not have had any matter in them, which could be governed by these laws, so an empty world (v), is a possibility for both of them, and so at that world it would be a nomic necessity that the inverse square law holds, and it would be a nomic necessity that the inverse cube law holds. But if world v is a (nomic) possibility for w1 and it is a (nomic) possibility for w2, then w1 as well as w2 should be (nomic possibilities) for v. But if both are a nomic possibility for v, then w1 should be a nomic possibility for w2 and vice versa,

---

[186] Fine (2002, 2005)

[187] Lange (2009)

[188] A detailed discussion of Fine's points can be found in Leech (2015). The current discussion cannot be quite as detailed.

[189] Fine (2005) p. 243ff

which is impossible, as it is precisely an impossibility in w2 that the matter be governed by the inverse square law, and vice versa for w1. So relativization apparently leads to contradiction.

This argument, however, will not do. For it implicitly equivocates between different nomic possibilities and necessities[190]: the nomic necessity relative to the inverse square law, and the nomic necessity relative to the inverse cube law. These are different natural necessities, and the empty world, which is a world in which both laws (trivially) hold, cannot work as a bridge between the two. Both w1 and w2 are nomic possibilities for world v, but they are so in two different senses, one in the sense that the inverse square law is a nomic necessity, the other in the sense that the inverse cube law is a necessity. The fact that the two spaces of possible worlds in which these two different laws hold overlap at the empty world, does not give us any license to move from the one to the other. One may object that it seems strange from the standpoint of the empty world that it should somehow preclude the possibility of w2, and vice versa for w1. But there is in fact nothing strange about that at all: The empty world, in virtue of trivially verifying all natural laws, of course does not by itself tell us which of them we should hold fixed when considering what is possible given some of these laws. But we do not consider which truths to relativize to from the standpoint of the empty world, precisely because it does not give us any good guidance on which truths to hold fixed, and also because it would be quite boring. In any case, it would be a different kind of nomic necessity, and not the one we care about when considering nomic necessity relative to the laws which hold at w1 or w2.

---

[190] See Leech (2015) p. 14ff for a detailed analysis of Fine's argument and her diagnosis of why it doesn't work, which has inspired my own diagnosis.

To put the point slightly differently: By relativizing necessity to the laws of nature which hold at the actual world, one does not commit to these laws being laws from the point of view of all worlds in which they are true (they are not laws from the point of view of the empty world), one only commits to the fact that the law is true at these worlds. After all, there is more to being a law than just being a true generalization.[191]

With this alleged counterexample out of the way, there still remains a challenge about how well the choice of truths, which nomic necessity is relativized to, is motivated. This issue may have been in part what motivated the attempted counterexample in the first place.

Here is what Fine has to say about the matter: "The general problem is that a definition of natural necessity as a form of relative necessity will tend to make the necessity of the propositions with respect to which the necessity is relative a trivial or insubstantial matter; yet we are inclined to think that the necessity attaching to the laws and the like is not of this trivial sort."[192] Marc Lange similarly argues that the modal force of laws is not captured by a relativization-account: "If being "naturally necessary" were nothing more than being a consequence of the laws, then for the laws to boast merely of their natural necessity would leave us unimpressed."[193]

The problem seems to be one that attaches to any relative-necessity account: The truths we relativize metaphysical necessity to will be necessary in the relative sense in a somewhat

---

[191] Even David Lewis' Humean account of laws requires the true generalizations that are to count as laws to be part of a best system that explains the going-ons at the world. See Lewis (1973).

[192] Fine (2002) p. 247

[193] Lange (2009) p. 48

insubstantial way, because we can just make them necessary by taking them to be the truths we choose to relativize to. So the necessity bestowed on the natural laws by the relativization account seems to be just a matter of our choice. I could, for example, think up a new kind of necessity relative to the truth that there is a cup of coffee on my desk, in which case it would be coffee-cup-on-the-desk-necessary that there is a cup of coffee on my desk.

I do admit that there is a sense in which it is true that the necessity of the laws of nature is a relatively insubstantial matter on the present account of nomic necessity. But in the sense in which I do admit it, it is quite plausibly so. In a different sense the necessity of the laws of nature should not be a trivial matter, but in that sense the laws of nature are not a trivial matter even on the relativist account.

The sense in which the necessity of the laws of nature should not be an insubstantial matter, is the sense in which it is an insubstantial necessity that it is necessary that the coffee cup is on my table, relative to the truth that the cup is on my table. One can, in this way, make up all kinds of necessities and make necessary any truth by just relativizing the necessity to these truths. So what is necessary in all these different senses, (coffee-on-the-table-necessity, raining-outside-necessity, etc.) is up to us, and as such not a substantial matter: we just make it up by relativizing appropriately. Clearly, nomic necessity should not be up to us in that way.

Luckily it isn't. For what turns out to be a law of nature is not up to us in any sense. What the laws of nature really are, is a substantial question, it is independent of what we make up, and we invest quite a bit of time and resources trying to find out about

it. So it seems the reasons for choosing the laws of nature to be the truths to relativize nomic necessity to, is not trivial after all. If the similarity between cases of stipulated relativization and relativization to natural laws was a worry, one can be assured that there is a substantial difference between the cases.

It must be admitted, however, that laws of nature are not laws of nature in virtue of having a special modal force, if natural necessity is to be explained via relativization. Rather, the direction of explanation is the other way around: The laws of nature have their modal force, that is, are nomically necessary, in virtue of being laws of nature. This, however, can hardly be taken to be a problem for the relativization strategy, as there are well established accounts of laws of nature which do not make reference to a special kind of modal force in explaining what laws of nature are.

The best known account of laws along these lines is the best system account.[194] On this account, a universally quantified proposition is a law of nature at a world, if it is both true, and a part of the simplest and strongest, for short, the best system systematizing all physical going-ons at a world.[195] This account does not make use of a modal criterion to individuate laws and consequently the direction of explanation advocated by the relativization account would be preserved.

Another account of laws which is also compatible with the proposed direction of explanation is, despite its name, Armstrong's necessitation account of laws.[196] According to this account a statement of the form "all Fs are Gs" is a law just in

---

[194] Most famously defended by Lewis (1973).
[195] Lewis (1973) p.73.
[196] Armstrong (1983)

case being an F necessitates being a G, where necessitation is a relation between universals, which explains the universal generalization. In contrast to the regularity view, it is not the regularity and some pragmatic considerations that make a law a law, but the fact that the regularity is accounted for by the necessitation relation holding between universals. That this necessitation relation holds between universals is itself a contingent matter, and so the laws are not themselves metaphysically necessary.

Lewis' and Armstrong's account, while well known accounts of laws of nature, are not the only theoretical options for explaining what natural laws are while avoiding to make use of a special kind of necessity. The proponent of a relativization account will have to endorse some such option.

It is worth noting, that theorists who have stressed that laws of nature should have a stronger connection to the modal realm, have often argued that laws of nature are metaphysically necessary.[197] On such an account, there would be no difference between metaphysical and nomic necessity, there would just be more metaphysical necessities than we expected. I mentioned above that this may be the case, but that I take it to be a fairly implausible view of what is involved in the reference realizing properties of the representations occurring in law-statements. Since such a view eradicates the distinction between metaphysical and nomic necessity it would still be compatible with the relativization strategy.

A philosopher who seeks to occupy a middle ground between those taking laws of nature to be accidental and necessary only in

[197] Shoemaker (1998)

so far as they are the basis of relativization, and those taking the laws of nature to be metaphysically necessary is Marc Lange. He proposes to think of laws as distinguished from mere accidents by a particular modal property: "maximal persistence under counterfactual suppositions"[198] What it is to be a law is to be stable in this sense. Lange thinks that there are contextual variations which explain an ordering of weaker to stronger, but relative to such a context, the maximally persistent propositions are genuinely necessary in virtue of being counterfactually stable. The beauty of this account is that it seems to unify two apparently incompatible demands on laws: On the one hand laws are supposed to be accidental, on the other, they are supposed to be necessary. Maximal stability in varying contexts explains how both can be the case.

However, it is not quite clear that the proponent of a relativization strategy cannot adopt a version of Lange's criterion for individuating laws, but deny that he has identified the direction of explanation correctly. For Lange thinks that counterfactual stability explains necessity, and this in turn explains lawhood. But one may reasonably cut out the middle man: counterfactual stability may be the epistemic mark of lawhood, but the necessity of laws may then be explained by them serving as the basis of relativization. This way of explaining the connection between counterfactual stability, lawhood, and necessity seems plausible, because it allows for an explanation of the counterfactual stability: The laws are stable, because they are the background according to which we evaluate counterfactual scenarios. They

[198] Lange (2009) p.90

serve as constraining conditions which we *take to be* necessary for the purpose of counterfactual reasoning.

These remarks suggest an answer the proponent of a relativization strategy may give to Lange, which takes seriously the observations about a connection between counterfactual stability and laws, but denies that this is what is definitive of lawhood.[199] I take it that this shows that monism about necessity remains a viable position in the face of Lange's criticism.

Summing up

In the above discussion I have defended monism about the varieties of necessity: the position that the source of the varieties of (well behaved) alethic necessity is metaphysical necessity. Weaker kinds of alethic necessity can be explained by relativization to certain other truths. Stronger kinds can be picked out by restriction to a subset of the metaphysically necessary truths. Two challenges to the relativization-account were considered. First an apparent counterexample was considered which could be shown to be ineffective. Second, I discussed the more serious challenge that laws are laws in virtue of being necessary, and not vice versa. I argued that while this challenge should be taken seriously, there is good reason for the modal monist to hold on to her position.

---

[199] A general argumentative strategy for the monist can be made out here: take whatever the proponent of the genuine necessity of laws proposes as explaining the special natural necessity of the laws which in turn is supposed to explain their lawhood, cut out the middle man, and take the feature to directly explain what laws are. Then explain the necessity by relativization. The effectiveness of this strategy depends on how plausible the identified feature is as an explanation of necessity or lawhood respectively.

I want to use this last paragraph to consider the implications of the discussion on the overall project of giving an explanation of necessity. The chapter attempted to show that the proposed explanation of necessity does not only give an explanation of one variety of necessity, but that the explanation can serve as an explanation of the source of other varieties of necessity, for example, nomic necessity. The strategy used to argue for this was rather orthodox, and involved taking the modal monist's perspective, and defining weaker kinds of necessity by relativization and stronger kinds by restriction. The proposed explanation of metaphysical necessity, however, provides us with a principled starting point for relativizing and restricting: It is the strongest notion of necessity which still provides for an intelligible notion of possible truth and it is the weakest notion such that it provides for everything that is, in the widest intelligible sense, possibly true. So there is some good reason to be confident that the proposed explanation of metaphysical necessity can be used as the basis for an explanation of a range of other varieties of alethic necessity. This would give the theory, apart from its potential as an explanation for metaphysical necessity, considerable further explanatory power.

A challenge to this may be mounted by those who take the varieties of necessity to have different sources. I examined this position criticized it on the grounds that it would detract from the generality and thus from the explanatory power of the proposed explanation of necessity. However, it would in principle be compatible with the present theory of metaphysical necessity that nomic necessity has a different source. There may be more than one explanation for why some truth cannot fail to be true. So while my official position is monism about the varieties of well-

behaved alethic necessity, nothing in the theory itself commits me to monism. I take monism to be independently plausible, however.

Before moving on to the discussion of essentialism and its interesting connection to the present theory, a short note on epistemic necessity is in order. Epistemic necessity is commonly characterized as that which has to be the case, given what we know; and epistemic possibility as that which we cannot exclude on the basis of what we know.

I said above that the strategy of restriction and relativization only works in cases of well-behaved alethic necessities. Epistemic necessity is not such a well-behaved alethic necessity, because it includes some truths as necessary which are metaphysically not necessary (the known contingent truths), and it excludes some metaphysical necessities as not necessary, because we do not know every metaphysical necessity. So the correlative possibilities, the truths we cannot exclude on the basis of what we know, will include some metaphysical impossibilities. Here is an example: it may be that we lack knowledge of the planets of our solar system and that for all we know Hesperus is not Phosphorus. From the given epistemic standpoint, this seems like a genuine possibility, but it is not, because it is metaphysically impossible. So epistemic possibility does not align nicely with metaphysical possibility.

Here is a suggestion on how this epistemic necessity may nonetheless be captured:

A proposition p is epistemically necessary for thinker a, iff p is known by thinker a, or it is known to be implied by what is known (by thinker a). Epistemic possibility is then defined as the dual of epistemic necessity: A proposition p is epistemically

possible, iff it is not epistemically necessary that it is not the case that p.

This formulation is sensitive to what a thinker knows to be implied by what he or she believes, and so it leaves open the possibility that a speaker is ignorant of the consequences of his or her beliefs as well as being ignorant about the modal status of what he or she knows. This account would divorce epistemic necessity and possibility from metaphysical necessity and possibility, as the account is relative to the epistemic capacities of a thinker. It can only be a suggestion, however, for a thorough treatment of epistemic modalities would require more space than can be given here. For present purposes, I merely suggest that we should take epistemic modality to be largely independent of metaphysical modality.[200]

[200] Modal rationalists like Chalmers would likely disagree, but I must leave this discussion for a different time.

Chapter 8: Essence and Representation

Essentialists take necessity to be explained by the essence of things, by those properties of objects which the object in question could not loose without ceasing to be the object it is. This essentialist explanation of necessity has some striking similarities to the present proposal and so a comparison between the two will be fruitful.

The connection between meaning and essence and their respective connection to definitions and necessary truth has been noted by eminent philosophers. W.V.O. Quine, a harsh critic of both essence and meaning, for example, states the connection between meaning and essence as follows: "Meaning is what essence becomes when it is divorced from the object of reference and wedded to the word."[201] And Kit Fine, a proponent of an essentialist theory of necessity, notes that we may understand what essences are in terms of giving a real definition of what an object is. But according to Fine, defining a *term* and giving a real definition of an *object* "is not merely parallel but […], at bottom the same."[202]

The present explanation of necessary truth in terms of reference determination does not trade in meanings, but focuses on what it takes for reference to occur, and it is not interested in giving definitions, but rather in what objects have to be like to be the referents of representations. But even with this new understanding

---

[201] Quine (1951a) p. 22. Associating the notion of meaning with essences, is also a rhetorical move by Quine to move meaning into disreputable company. His skepticism towards both meaning and essences has not stood to changing philosophical fashion, however, and so his remarks are here cited for their content only.

[202] Fine (1994) p.13.

of analyticity, the similarities between the notions are quite obvious.

I want to make an even stronger point, however. I will argue that the present theory can be viewed as giving an essentialist account of necessity in virtue of explaining what essences are. By doing so, it also provides us with an explanation of how essence gives rise to necessity, an explanation orthodox essentialist proposals lack.

To convince you of this strong connection, I will first introduce the orthodox essentialist explanation of necessity. Second, I argue that such essentialist accounts have a shortcoming, because they fail to make clear how the necessary truths are true in virtue of the essence of the objects they are about. Third, I will show how the present theory can fix this by both giving an account of what essences are as well as explaining how essence gives rise to necessity.

1. Essence and Modality

Essence, Kit Fine forcefully argued, cannot be explained by necessity.[203] Rather, he suggests, necessity should be explained by essence: "It seems to me that far from viewing essence as a special case of metaphysical necessity, we should view metaphysical necessity as a special case of essence. For each class of objects, be they concepts or individuals or entities of some other kind, will give rise to its own domain of necessary truths,

[203] The counterexamples to a modalist account of essence is by now well-known and so I will not discuss it here in detail. The idea of the counterexample is that while some properties may be necessarily had by an object, these properties are intuitively not essential to the object (for example being such that 2+2=4), and so it is not sufficient for being an essential property of an object that the property is a necessary property of the object.

the truths which flow from the nature of the objects in question. The metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatever"[204] So according to Kit Fine, the metaphysically necessary truths are the ones which are true in virtue of essence, those which flow form the nature of things.

This explanation of necessity by essence is not without its attractions, for it apparently provides us with an explanation of necessity that is not in need of postulating possible worlds, be they abstract or concrete. Also, the explanation does not have to be primitivist about some modal notion, and it uses a philosophical notion now widely accepted as interesting in its own right, that is, essence. Further, if essence indeed cannot be explained by necessity, then there seems to be little room for explaining the apparent connection between the two, except by taking essence to explain necessity.[205] Accordingly, the idea was taken up by a number of philosophers: Bob Hale, for example, writes: "It is necessary that *p* because it is true in virtue of the natures of $X_1, \ldots, X_n$ that *p*"[206], and Jonathan Lowe equally weighs in: "essences are the ground of all metaphysical necessity and possibility".[207]

This attractive essentialist explanation of necessary truth leaves open two questions, however. First, if essence is not to be explained by modality, what is it to be an essential property?

---

[204] Fine (1994) p. 9.

[205] There are further options apart form one explaining the other, or vice versa. Either the apparent connection between the two could be a mere coincidence, but, much more interestingly, a third element may explain both. The explanation of modality and essence outlined in what follows can be understood as being an explanation of the connection between the two by a third element.

[206] Hale (2013) p. 145

[207] Lowe (2008a) p. 23

Second, what is truth in virtue of essence supposed to be? While there may be no urgent need to give an answer to the first question, as explanation has to stop somewhere and we may already have sufficient grasp of what essentiality consists in through the notion of real definition, the second question is more urgent, for it concerns the core explanatory claim of the theory and it is not obvious that the essentialist theory, as it stands, has an answer to it. So let me move straight to the core explanatory claim of essentialist theories of necessity.

The essentialist idea is that necessary truths like the one expressed by "Socrates is human" are necessarily true (if Socrates exists), because it is true in virtue of it being essential to Socrates that he is human. This surely has a nice ring to it, but on closer inspection it looks slightly mysterious how the 'because' in this explanation is justified. For "Socrates is human" does not seem true in any special way. It is just true in virtue of Socrates being human. The fact that being human is essential to Socrates adds nothing to the explanation of the truth of the proposition. So it seems that the fact that being human is an essential property of Socrates is not involved in an explanation of the truth of "Socrates is human" rather, it is the mere fact that Socrates is human that is involved in the explanation of the truth of "Socrates is human". What role the fact that being human is essential to Socrates plays in the explanation of the truth remains unclear. To get this challenge into focus, I will first compare it to a related objection to theories utilizing the notion of truth in virtue of meaning, before making a seemingly natural proposal for a solution, and showing why this solution doesn't work.

The problem is reminiscent of an objection commonly leveled against the notion of truth in virtue of meaning, and there widely

accepted as decisive, which says that there is no such thing as truth in virtue of meaning, because it is, in the appropriate sense, the world which makes a sentence true, not its meaning. All meaning does, is tell us what a sentence says, but what makes it true is that what it says really is the case. Consequently there is no such thing as truth in virtue of meaning.[208]

A similar argument seems warranted in the case of truth in virtue of essence: The fact that some property of an object is an essential property of it is irrelevant to the explanation of the truth saying of the object that it has this property. Whether it is true that Socrates is human doesn't depend on him being essentially human, it merely depends on him being human. So there is no such thing as being true in virtue of essence, provided "truth in virtue of essence" is a special way of being true, different from regular ways of being true. Here Quine's observation that meaning and essence are two faces of the same coin, comes back to haunt the essentialist.

Maybe, however, the essentialist has a way to answer the objection the proponent of truth in virtue of meaning doesn't have. Maybe the suggestion is weaker than I interpreted it. Maybe it is not that "Socrates is human" has a special way of being true, maybe being true in virtue of the essence of something just means that the proposition is true in virtue of Socrates being human, plus the fact that being human happens to be an essential property. This suggestion does not exhibit the problem just outlined, for it does not claim that there is a special way of being true. It is just

---

[208] As was outlined in some detail above, Boghossian (1996) gives the now standard articulation of this worry. But he is neither the first nor the last to use it to argue against truth in virtue of meaning. See for example Lewy (1976), Williamson (2006). For the most famous criticism of analyticity as truth in virtue of meaning, see Quine (1951).

that the properties of Socrates which make the proposition true also happen to be among his essential properties. So we may separate the facts that make true the proposition (Socrates' being human), from the facts that make it necessary (Socrates being essentially human). This seemingly natural solution, however, is problematic, because it does not explain how the fact that being human is an essential property of Socrates gives rise to the necessity of the proposition. If truth in virtue of essence was a special way of being true, this would at least suggest that this special way of being true gives rise to a special mode of truth, but since there is no special way of being true, no hint is given as to how the status of being an essential property gives rise to necessary truth.

Note that I am not denying that if it is essential to Socrates that he is human, then the proposition expressed by Socrates is human is necessarily true. We may well be able to pair every necessary truth with some essentialist truth and find them to be logically equivalent. The problem is that an explanation of how essence gives rise to necessity is lacking, and thus the account lacks in intelligibility.

If standard essentialist proposals do not properly explain how essence gives rise to necessity, how are we to understand the claim that essence gives rise to necessary truth? I suggest that the essentialist should adopt the explanation of necessity given by the present theory, and think of essences as the reference realizing properties determined by singular representations. This will allow for an illuminating reading of the slogan 'truth in virtue of essence'.

2. Essences as Reference Realizing Properties

With the help of reference realizing properties, we can give an account of what the essential properties of something are: The essential properties of an object are those of its properties which realize, or would realize, the reference relation between the object and a singular representation referring to it. This identification of essences with reference realizing properties allows for a clarification of the essentialist explanation of necessary truth.

That the identification is warranted can best be seen when considering how we may characterize essential properties as well as reference realizing properties. The essential properties of an object can be characterized as properties an object could not loose without ceasing to be the object it is.[209] Reference realizing properties can similarly be characterized as properties the referent of a singular representation could not loose without ceasing to be the referent of the representation. So provided that Socrates is the referent of the representation ⌜Socrates⌝, if and only if Socrates exists, both the essential properties of Socrates and his reference realizing properties are properties he could not loose without ceasing to be Socrates. That essential properties and reference realizing properties are both properties an object could not loose without ceasing to exist suggests that an identification is warranted, even if the epistemic possibility remains open that the essential properties differ from the reference realizing properties, and are just (necessarily) co-instantiated.

---

[209] This characterization, as well as related ones (the properties that constitute an object's identity, the nature of an objecte, etc.) are commonly not intended as analyses of essence, nor as constitutive explanations, but merely as elucidations that help us understand, and pick out what essences are.

I will try to argue against remaining doubts by pointing out the explanatory advantages of the identification. The first and foremost advantage is that the identification provides us with a genuine explanation of necessary truth in terms of essences. If the reference realizing properties of objects are their essential properties, then the explanation of necessity can proceed along the lines suggested above. Assume, for example, that it is essential to Socrates that he is human and that it is therefore a reference realizing property of Socrates, determined by the singular representation ⌜Socrates⌝, that he is human. Given this, it is necessarily true that Socrates is human, because the requirements for truth of the truth-evaluable representation ⌜Socrates is human⌝ are guaranteed to be satisfied. This is the sense in which necessity flows from essence: The essence of Socrates, the relevant reference realizing properties of Socrates, explain why the requirements for truth of the representation ⌜Socrates is human⌝ are guaranteed to be satisfied.

In contrast to the proposal of how essence and modality may be connected according to the orthodox essentialist which just noted that an essential property was involved in the truth-making, a genuine explanation of how this guarantees the truth of the representation can be given along the lines of the present theory: A representation cannot fail to be true in virtue of the appropriate relationship between requirements for truth and reference realizing properties.

One further advantage of taking essences to be reference realizing properties should be mentioned here. Essentialists occasionally

say things like "essence precedes existence"[210]. What is meant by this, is that the fact that some object has a certain essence is not conditional on its existence, even though the instantiation of the essence is a condition on the existence of the object. This is supposed to allow for the possibility of non-existent objects having essences as well.[211] But this talk of essence preceding existence is slightly mysterious. In what sense does something have an essence without existing? Identifying essential properties with reference realizing properties explains in which sense something can have an essence without existing: Since what the reference realizing properties of a singular representation are, is determined by the reference determining properties of a representation, there is no need for an object actually having these properties for there to be a fact of the matter what properties such a referent would have, if it was the referent of that representation.[212] So while the reference realizing properties are not instantiated, it is nonetheless determined what they would be if the relevant object existed. Thus there is good sense to be made of essence preceding existence.

The explanatory advantages of moving from essences to reference realizing properties, and preserving some central essentialist tenets without incurring the problems of more orthodox essentialist proposals, may not yet have convinced the committed essentialist. For the essentialist may take the present proposal to

---

[210] Lowe (2008a)

[211] It is not mandatory for the essentialist to insist that non-existent objects have essences. But to be able to account for necessary truths about non-existent objects, there seems to be some pressure to do so.

[212] As was discussed above, there may be some doubts that singular reference is possible without the existence of an object. This is an open question, which should not impede on the present point, however.

be so far removed in spirit from his own project that an identification is too much of a stretch. If the reader should be of such a conviction, I merely ask her to acknowledge the similarity to the essentialist proposal as well as the explanatory merits of the explanation of necessity in terms of reference realizing properties.

Summing up

Comparing essentialism to the present theory revealed that the theories are closely related. I have tried to convince you that one may even take the present theory to be an improvement on the orthodox essentialist theory, because it both offers an explanation of what essences are, and of how essences explain necessary truth. Essences may be identified with the reference realizing properties singular representations determine.

The connection between meaning and essence has been noted before and Quine rightly identified the two notions as two faces of the same coin. So it is not surprising that the essentialist theory of necessity encounters problems similar to those of traditional accounts of necessary truth in terms of meaning. Traditional accounts of necessity in terms of meaning sought to explain necessity by identifying it with truth in virtue of meaning. This view is untenable, because necessary truths just are not true in virtue of meaning. But truth in virtue of essence is also mysterious, because it is unclear what it is to be true in virtue of essence. So the move from the word to the world, while understandable, does not, by itself, succeed in giving a satisfactory explanation of necessary truth.

Essentialists are right in looking to the world for an explanation of necessary truth, and away from the words and their meanings. But by looking for essences in the objects, independently of the

representations, essentialism subscribes to a problematic view of the relationship between representation and world as two distinct realms: the realm of representation, and the realm of the world. This sharp distinction was already the reason why an explanation of truth in virtue of meaning failed. Moving the explanation of necessary truth from the realm of words and meanings to the realm of objects and essences does not succeed, because what is needed is an explanation of how representations and what they represent work together to produce reference, truth, and the modes of truth. This synthesis is made possible by the proposed framework which allows for the identification of essences with reference realizing properties.

Chapter 9: Evaluating the Explanation of Necessity

The previous chapters developed and defended the theory of necessity which I take to give a good explanation of metaphysical necessity. I want to use this chapter to summarize and emphasize the points which speak in the theory's favor. Here is the main reason for believing the theory to be a good theory of necessity: The theory offers a non-circular, illuminating and intelligible, as well as general explanation of metaphysical necessity that appears extensionally adequate, accounts for puzzles about possibilia and necessary existence, explains what essences are and along the way makes only minimal assumptions about the nature of representation and its connection to the world. I take these to be reasons for taking the explanation to be a good explanation, but let me go over each point in turn.

1. Non-Circularity

A philosophical explanation, I argued in the first part, should give a non-circular, intelligible, and general constitutive explanation of a philosophical phenomenon. That is, it should tell us what it takes to be X, where X is the phenomenon to be explained. Giving such an explanation may involve conceptual analysis, but need not, as there just may not be the conceptual or linguistic resources to analyse a concept into simpler parts. The case of an explanation of necessity is such a case. No (credible) explanatory analysis is available which illuminates what it takes for a

representation to be necessarily true.[213] Nonetheless, a constitutive explanation of why a representation is necessarily true can be given. Such an explanation should, by identifying what it takes for a representation to be necessarily true, make clear what gives rise to necessary truth. A central condition on this explanation is non-circularity: The explanation should not again involve what is in need of explanation.[214] The official explanation of necessary truth I proposed was the following:

> *Explanation of Necessity:* A necessarily true truth-evaluable representation is necessary, because the reference determining properties of the constituent referential representations determine the reference realizing properties such that the requirements for truth of the truth-evaluable representation are guaranteed to be satisfied.

I already argued above that the proposed constitutive explanation of necessity is not circular. So let me rehearse this reasoning,

---

[213] The proponent of possible worlds would argue that an illuminating analysis is available, namely an analysis in terms of possible worlds. However, even granting the biconditionals expressing the analysis, it is not clear that this helps with philosophical explanation, for it must also be true for the analysis to contribute to an explanation that it makes clear what it takes for a representation to be necessarily true. If this is assumed for the analysis in terms of possible worlds, one is committed to the existence of a multiverse of possible worlds. A commitment which can be, and should be avoided.

[214] We must take care not to confuse this non-circularity requirement on constitutive explanation with the much more specific non-circularity requirement on conceptual analysis, for the two have a different rationale: A conceptual analysis cannot be true if it involves the analysandum in the analysans, for nothing can be both a proper constituent of a concept and the concept itself. Arguably this kind of circularity is quite rare, however, and usually the circularity charge against conceptual analyses is that the analysans involves some concept from the same 'family' as the analysandum and is therefore not properly illuminating.

because the official explanation may at first glance appear circular, for it seems to say that a truth-evaluable representation is necessarily true, because it is guaranteed to be true. This would be rather unilluminating, for isn't a guarantee for truth just necessary truth?

Part of the charge should be admitted: being necessarily true just is being guaranteed to be true, and this by itself is unilluminating. However, the explanation also says how this guarantee for truth comes about, it explains what it takes for the guarantee of truth to be effected: The reference realizing properties determined by the constituent referential representations must stand in the appropriate relationship to each other and to the requirements for truth of the truth-evaluable representations. What this "appropriate relationship" is, depends on the requirements for truth of the truth-evaluable representation and so in order to state the explanation in a general way, this relation is picked out by its effect, that is, the guarantee for truth. Nonetheless, the appropriate relationship, which is ultimately constitutive of necessary truth, can be given for each (kind of) representation. The appropriate relationship between the reference realizing properties determined by two referential representations which require for their truth that the referents of one representation are among the referents of the other, for example, is that the reference realizing properties of the other are among the reference realizing properties of the one. Similar accounts can be given for other kinds of representations.

I conclude that no circularity is involved in the present explanation of necessary truth. It is constitutive of necessary truth that there is an appropriate relationship between the reference realizing properties determined by the constituent referential representations and the requirements for truth of the truth-

evaluable representation. What this appropriate relationship is, can be specified for each necessary truth without using modal notions.

2. Intelligibility

Intelligibility is the second criterion for a good philosophical explanation. It is admittedly a somewhat loose criterion, but there is nonetheless some sense to be made of it, and it should be recognizable, if an explanation meets it.

Explanation is supposed to move us closer to an understanding of the phenomenon which is explained. It should help us to see how a certain phenomenon came or comes about.[215] This seems to hold both for genealogical and constitutive explanations.

Take, for example, a causal explanations: if we want to know why the light is on (causally), we want to be told something that makes us see how it came about that the light came on. The relevant fact which explains this in the present case, may be that someone switched on the lights. Together with our background knowledge that there is some appropriate causal mechanism connecting the switch and the light, this information allows us to see why the light is on.

Constitutive explanations are not very different in this respect. While we do not want to be told about some causal mechanisms, we want to be told about the metaphysical ‚mechanism'. We want to be told that what is constitutive of a phenomenon is in place. If we want to know, for example, why (constitutively) there is a triangle on the sheet of paper, an explanation would list the constitutive features of being a triangle and say that they are in

[215] 'Comes about' should here not only be read as causally becoming, but also as constitutively coming to be.

place: there is a triangle, because there is a closed figure with three sides on the sheet of paper.

Since what we are looking for in philosophical explanations are constitutive explanations, what is needed from an intelligible explanation is that it helps us to see how the phenomenon to be explained comes to be. In the special case of necessary truth it should therefore allow us to see how it comes about that a necessary truth cannot fail to be true. The present explanation of necessity does just that. It describes the mechanism by which some representations cannot fail to be true: They cannot fail to be true, because necessarily true representations are so constituted that the relationship between the referents of the constituent representations required for its truth is bound to obtain in virtue of how reference of the constituents is determined. So we can see how the truth of a representation is guaranteed.

Given the outlined (rather minimal) assumptions about truth and reference, it is no surprise any more that necessary truth arises. Some representations are guaranteed to be true in virtue of how reference determination and truth work together.

One way to test the intelligibility of an explanation is to see whether the explanation provides us with a recipe for bringing about the phenomenon we want to explain. In the causal case, a good explanation allows us to reproduce the effect we wanted to explain, in the constitutive case it allows us to recreate the phenomenon we intend to explain.

The given explanation of necessary truth passes this test. It provides us with a recipe for making representations that are necessarily true: We have to combine referential representations with adequately related reference realizing properties which they determine to form truth-evaluable representations with

requirements for truth that are guaranteed to be satisfied by the relationship holding between the reference realizing properties.
Thus I conclude that the theory also satisfies the desideratum of intelligibility. It allows us to understand how necessary truth comes to be.

3. Generality
The third criterion which is supposed to tell us something about the theory's quality is generality. Generality is a requirement for philosophical theories, because it is one of the tasks of philosophy to give general constitutive explanations of phenomena. So an explanation of necessary truth should explain how necessary truth arises generally and not just for a few isolated cases.
As I mentioned in the first part, the criterion is a little elusive, because it applies only in so far as the phenomenon to be explained really is uniform: A phenomenon which at first appears uniform may turn out to require many different explanations and so it may turn out that generality is a misguided requirement on some explanations. Nonetheless, if a satisfactory general explanation can be given, its generality speaks in its favor, for it allows us to understand a wide range of cases, and holds true to the appearance that the phenomenon in question really is uniform.
So is the present explanation of necessary truth a general explanation of necessary truth? I believe that the argument above shows that it is general in the appropriate sense, for the explanation is in principle applicable to every necessarily true truth-evaluable representation. The fact that it seems to capture standard examples of necessary truths in a uniform way is reason to believe that it is likely to apply to the other relevant cases. I thus think it is safe to say that the present theory is general in the

intended sense, for it makes sense of the phenomenon of necessary truth quite generally.

4. Extensional Adequacy

Extensional adequacy is a minimal requirement for successful philosophical theories. The requirement itself derives its force from being a strict requirement on conceptual analyses. There the requirement is that the biconditional used to state the analysis must be true. If cases can be found where either the left-hand side or the right-hand side of the biconditional turn out to be false, while the other is true, then this demonstrates that analysans and analysandum represent two different phenomena. The present explanation, however, is not a conceptual analysis. So applying the standard of extensional adequacy to the present theory is not as straightforward. Nonetheless, we can make some sense of the requirement. For it should still be the case that the explanation only explains, that is, renders necessarily true, all and only those representations that are necessarily true.

If the explanation is the true explanation, this criterion is trivially satisfied, but to test whether it has claim to being the true explanation, we can see whether it renders the standard examples of metaphysically necessary truth necessarily true. If it doesn't, we are either wrong in our opinions about the clear cases, or, which is more likely, we are not explaining what we are trying to explain. So it is important that the theory covers the clear cases, the cases we take to be paradigms of metaphysical necessity.

These clear cases are what sometimes runs under the label of 'broadly logical truths', as well as 'essentialists truths'.[216]

There are three reasons to take the theory to be extensionally adequate. First, examples show that it explains the necessary truth of paradigmatic truths from all these domains. Second, the explanation of necessity yields a notion of necessity which is the strongest notion of necessity that still provides for an intelligible notion of possible truth, and the weakest notion such that it provides for everything that is, in the widest intelligible sense, possibly true. Third, elements of the explanation can be taken to give an account of what essences are, the explanans of a theory of necessity which is relatively widely accepted as being extensionally adequate. Let me go over these three reasons in slightly more detail. None of the arguments is quite decisive, but together I believe that there is strong reason to suppose that the theory is extensionally adequate, that is, that it really is an explanation of metaphysical necessity.

The first reason for taking the theory to explain metaphysical necessity is that it covers standard examples of necessary truths. In the previous chapters, examples of logical truths, conceptual truths, mathematical truths, and essence attributions were all shown to, at least in principle, be covered by the proposed explanation. This shows that there is no domain of standard examples of necessary truths which are not, at least in principle,

---

[216] The terminology is not quite uniform in the debate. For some, metaphysically necessary truths just are the broadly logical truths, for others the broadly logical truths are only conceptual and logical truths, and not essentialist truths. Be that as it may, however, a theory of metaphysical necessity should be able to account for the necessity of standard cases of conceptual truths, logical truths, and essentialist truths.

covered by the present account. There may be two weak points in this argument from examples.

First, a few examples may not sufficiently raise confidence in the claim that the explanation is applicable to all necessary truths. This first weak point, however, is itself quite weak, for given that the examples are paradigms of their respective domains, there seems to be little reason to believe that the explanation does not apply to further examples as well.

Second, the examples relied on assumptions of what the respective reference realizing properties determined by the referential representations are. These assumptions are integral to the success of the explanation, and so there may be doubts that these assumptions are sufficiently justified.

I relied on two ways of justifying the assumptions. First, I appealed to our understanding of the representations expressed by the example sentences and our intuitions about what it takes to be the referent of the constituent representations. I take it that the assumptions based on this were at least plausible. It may be that a more detailed linguistic analysis of the example sentences shows that the representations expressed by the sentences are different from the representations I suggested they express. So it may turn out that the representations commonly expressed by the example sentences are constituted by different referential representations which determine different reference realizing properties or that the requirements for truth turn out to be different. If so, then it may be that some of the examples do not turn out necessarily true on the present account. However, in this case we would have to ask ourselves whether our intuitions about the necessity of the standard examples did not rely on the same misunderstanding of the sentences as the present explanation. And so it may turn out

that while some of the example sentences do not commonly express a representation that is necessarily true according to the present theory, there is a representation expressed by a similar sentence which is indeed rendered necessarily true by the present theory.

In light of the further reasons for extensional adequacy, I trust that a more detailed linguistic understanding of which representations are expressed by sentences will help to validate, and to make more precise, the sketchy account of reference realizing properties and requirements for truth given in the examples-section above.

The second reason for taking the theory to be explaining genuine metaphysical necessity was discussed in the chapter on the varieties of necessity: The kind of necessity explained by the present theory is the strongest kind of necessity which still yields an intelligible corresponding notion of possible truth. This, I suggested, should be exactly what we would expect from metaphysical necessity and consequently there is good reason to take the present theory to be a theory of metaphysical necessity.

Let me quickly go over the argument again: Anything which is not necessarily not the case is possibly the case. So each variety of necessity has its own corresponding notion of possibility. However, if there is a notion of necessity stronger than the necessity explained by the present theory, the corresponding notions of possibility cannot be construed as being possibly true, for there are bound to be some allegedly possible representations that just cannot be true. Here is an example: Take the metaphysical impossibility that there are married bachelors. It is metaphysically impossible, because the requirements for truth of the representation representing that there are married bachelors cannot be satisfied in virtue of how the reference of the

constituent representations is determined. But that there are married bachelors is a logical possibility, because it is not a logical necessity that there are no married bachelors. However, since the requirements for truth cannot be satisfied, there is no sense to be made of the logically possible truth of a representation representing that there are married bachelors: the representation, if it is to be the representation it is, *cannot* be true in any intelligible sense of true.

This example is not an isolated case. Any kind of necessity stronger than the necessity explained by the present theory will yield a corresponding notion of possibility which cannot be interpreted as possible *truth*. The reason for this is that the explanation of necessity is directly connected to what is required for the truth of the representation. So anything that is not possible according to the present theory cannot be true simpliciter. Since this is just what metaphysical necessity is supposed to be, the strongest kind of necessity which still yields an intelligible notion of possible truth, the present theory should be taken to explain metaphysical necessity.

This reasoning relies on certain assumptions about what metaphysical necessity should be, namely that is the strongest notion yielding an intelligible notion of possibility. This may be challenged. I take it, however, that the explanation of necessity captures two quite intuitive aspects of metaphysical necessity which are present in the criterion given: First, that what is metaphysically necessary is true no matter what. Second, that what is metaphysically possible is possibly true in the widest possible sense.

The third reason for taking the explanation to account for metaphysical necessity is that it can be viewed as underwriting an

essentialist theory of necessity which is commonly taken to be extensionally adequate. I claimed that essences can be identified with the reference realizing properties determined by singular representations. Why? Because the reference realizing properties are those properties an object has, if it is the referent of the singular representation referring to it, so given that a singular representation refers to an object, if and only if the object exists, the object has these reference realizing properties, if and only if it exists. This also has a modal dimension: If an object would loose the reference realizing properties, then it would no longer be the referent of the singular representation, and given that the singular representation refers to the object as long as it is the object it is, it would no longer be the object it is. This characterization of reference realizing properties is parallel to the characterization of essential properties as those properties which an object could not loose without ceasing to be the object it is. So while the essential and the reference realizing properties may be two distinct or merely partially overlapping sets of properties, which are both a subset of the properties an object necessarily has (if it exists), it seems that the obvious way to go is to identify the essential properties with the reference realizing properties, especially given that an explanation of what essences are and how they give rise to necessity is needed.

These three reasons speak in favor of taking the theory to really be a theory of metaphysical necessity, a theory that explains why the truths that are metaphysically necessary cannot fail to be true. None of these reasons is a knockdown argument, and all depend to a certain extend on assumptions which may be challenged. All three together, however should raise the confidence in the theory's extensional adequacy.

5. Puzzles Solved

Bertrand Russell famously claimed that "a logical theory may be tested by its capacity for dealing with puzzles"[217]. And so I want to take his advice and bring to mind again the unified solutions the present theory offers to some puzzles in modal metaphysics, namely puzzles about necessary existence, mere possibilia, and impossible objects. If the proposed explanation of necessity is accepted, it offers a quite natural solution to these puzzles. Since it can account for the truth of talk about mere possibilia, it may also be viewed as a properly actualist theory of modality. This is a desirable feature, not least for allowing an answer to an apparent dilemma for an explanation of necessity posed by Simon Blackburn.[218]

Let me rehearse the puzzles and how the theory and the framework in which it is developed solves them. First, there is the problem of necessary existence: If it is necessarily true that Socrates is human, then it seems that it is equally true to say that necessarily, Socrates exists. For it seems to be a requirement on the truth of Socrates' being human that he exists. Socrates, however, is a contingent being. He does not exist necessarily. This leaves us with a puzzle, for it cannot be that Socrates both exists necessarily and contingently. The present theory offers a solution. For on the present theory it is not necessarily true that Socrates is human, provided the representation ⌜Socrates is human⌝ requires the existence of Socrates for its truth. It is, however, necessarily true on the present theory that Socrates is human, if the representation does not require for its truth that Socrates exists. So

---

[217] Russell (1905) p. 484
[218] Blackburn (1993)

the theory does not imply that Socrates necessarily exists, thereby solving the puzzle. That one should take the representation expressed by "Socrates is human" to not require for its truth the existence of Socrates, in so far as the sentence is supposed to express a necessary truth, is plausible on a very basic application of the principle of charity: One should, as far as possible, interpret someone as speaking the truth. Since it is only necessarily true, according to the present theory, that Socrates is human, if this doesn't require Socrates' existence for its truth, the sentence should be interpreted as expressing this representation.

The puzzles about mere possibilia and impossible objects are solvable once we accept that some representations do not require the existence of the object they are about for their truth. What is puzzling about mere possibilia, objects that do not actually exist, but may have existed and impossible objects, objects that do not exist and could not have existed, is that there are some (necessary) truths about them which seem true, but at the same time the truth of such representations seems to require the existence of these very objects. Since it was acknowledged in the solution to the puzzle about necessary existence that some representations do not require for their truth that the object they are about exists, the existence of mere possibilia, or of impossible objects, is not required for the truth of some representations about them. Especially in the case of essence attributions to mere possibilia and impossible objects, no existence of the object is required for the truth of the representation. So there is sense to be made of talk of possibilia without having to admit that merely possible objects exist.

The ability of the theory and the framework in which it is developed to give a satisfactory solution to these puzzles is thus a

further reason for taking the proposed explanation of necessary truth seriously.

A further puzzle: Blackburn's dilemma

There is a further puzzle which has not been mentioned in the parts above, to which the present theory and its general framework allow a good and satisfying answer: Blackburn's dilemma. The dilemma is supposed to show that no explanation of necessary truth can succeed, for either the explanans is itself necessary, then it seems that there is a residual necessity itself in need of explanation, or the explanans is contingent, then it seems that it could have been different and thus that it could have been the case that the explained necessity was not necessary after all, a conclusion which undermines the necessity of the explanandum.[219]

It is not hard to get a sense that something must be wrong with this reasoning, but it is not easy to pinpoint exactly where it goes wrong. So let me look at the dilemma in slightly more detail, and propose a solution for it that shows that the modal status of the explanans is irrelevant to the explanation. This will fit nicely with the broadly actualist leanings of the present theory and its explanatory strategy. While the solution to this puzzle is inspired by the present explanation of necessity, it does not depend on it and can be used by philosophers who do not accept the proposed explanation. As such, the solution to this puzzle does not speak directly in favor of the present theory, but an answer is nonetheless desirable, especially since there have been some

[219] Blackburn (1993)

doubts that a theory along the suggested lines could answer the challenge posed by the dilemma.[220]

I will first start with the contingency horn of the dilemma, discuss possible answers, offer my solution, and argue that it also solves the necessity-horn of the dilemma.

Blackburn's dilemma - the contingency horn

The argument for the contingency horn can be reconstructed in the following way in quasi-logical notation[221]:

(1) $\Box p$	Explanandum
(2) $\Box p$ because q	Explanation of $\Box p$ by q
(3) $\neg q \ \Box\!\!\rightarrow \neg \Box p$	Assumption: Plausible counterfactual principle about explanation
(4) $\Diamond \neg q$	Assumption: Contingency of q
(5) $\Diamond \neg \Box p$	From (3) and (4)
(6) $\neg \Box \Box p$	From (5)
(7) $\Box \Box p$	From (1) by S4-Axiom ($\Box p \rightarrow \Box \Box p$), contradicting (6)

The argument is valid as it stands and it also seems, at first sight, intuitively plausible: Assume that some necessary truth p is explained by some contingent truth q. Since q is contingent, it could not have been the case that q. But if q had not been the case,

---

[220] compare for example Hale (2013), who uses the dilemma to argue for his brand of essentialism against a representation-based account.

[221] cf. Hale (2002), Hanks (2008), Lange (2008) and Morato (2014). Some authors take p and q to stand for propositions, others take them to directly stand for facts. Nothing turns on the issue in what follows.
'$\Box$' is 'necessarily, '$\Diamond$' is 'possibly', and '$\Box\!\!\rightarrow$' is the counterfactual conditional.

it would not have explained necessarily p. Therefore "it is possible that p" is not necessary. But if it is possible that p is not necessary, then it is not really necessary. Consequently, a contingent explanans cannot explain necessity. Note that the notion of explanation used here is quite general, and one may read 'p because q', as 'p is grounded in q'. Nothing in the argument and general reasoning depends on any specific conception of explanation.

As has been pointed out by Hale and other authors following him, the contingentist seems to have two options to counter this argument.[222] First, she could reject the counterfactual dependence of the explanandum on the explanans, second she could reject the step from 1 to 7 above by rejecting the so called S4-Axiom, that what is necessary is necessarily so. The second option appears to be the most straightforward solution, for it just seems to beg the question against the contingentist's proposal that one should assume that what is necessary is necessarily so.

However, there is independent reason to believe that the strategy of rejecting S4 is not helpful, for, as Hanks points out, there is a different version of the argument which does not make use of the allegedly controversial principle, but instead assumes that the explanation takes a different form.[223] An explanation of necessity in terms of a contingent explanans may take the following form: □p because (p because q). Informally, that may be understood as stating that p's necessity is explained by the explanation of p in terms of q. Together with the plausible counterfactual principle about explanation, that if q would not have been the case, then p

---

[222] Hale (2013)

[223] See Hanks (2008) for an argument to this conclusion from which the following is a summary.

would not have been the case, it is easily derived that p is not necessary.[224] Here is how the argument goes: Since q may not have been the case, it follows by counterfactual dependence of the explanandum on the explanans that p may possibly not have been the case. But if possibly not-p, then it is not necessary that p. This follows without any appeal to the allegedly controversial S4 principle.

This argument indicates that what is at issue in the present case is not the correct modal logic, but rather the correct form an explanation of necessity should take, as well as the principles it should follow. That S4 is a red herring can be seen best, when we translate the above arguments into talk of possible worlds. This will also indicate how an explanation of necessity in terms of something contingent should be understood.

The right form of explanation

The proponent of the contingency-horn of Blackburn's dilemma thinks of an explanation of necessity in terms of contingencies in the following way: That something is true in all possible worlds is explained by contingent features in each possible world.

Take first the form of explanation Hanks presupposes for his argument: p is true in all possible worlds because (p because q) is true in each possible world. Formulated in terms of possible worlds, it is immediately obvious that such an explanation of necessity can never work, if the explanans is contingent. For if the explanans, q, is contingent, then there is a possible world in which it is not the case that q, but if there is a possible world in which it is not the case that q, then, given the factivity of explanation, it is

[224] Note that it would also be very implausible to reject the counterfactual principle about explanation in the present case.

also not the case in that world that p because q. Therefore, it is not the case that in every possible world it is the case that p (provided, that q's obtaining is a necessary condition for p's obtaining), but then p is just not necessary.

The same holds for the simpler form of explanation □p because q, where q is again contingent. In terms of possible worlds, this comes down to: in every possible world p, because in each world q. Which of course cannot be a good explanation, because q is not true in every possible world in virtue of being contingent. So how could it possibly explain p in worlds in which it is not true? Obviously it can't, and therefore this explanation cannot work.[225]

Both explanations, no matter what we fill in as p and q, can be seen to not get off the ground once they are spelled out like this, but there must be something wrong with any reconstruction of the contingentist explanation that attributes such an obviously wrong thesis to a proponent of this kind of explanation. It seems that this must be an extremely uncharitable reading and may indicate that the opponent is merely a straw man.[226]

The form the contingentist's explanation should take

So what could an explanation of necessity in terms of a contingent feature of the world look like instead? The obvious solution is to explain the necessity of p via the obtaining of q in the actual

[225] Lange (2008) suggest just such argument against the contingentist. (p. 125-126)

[226] It must be admitted that the criticism is not directed against a straw-man, if the opponent is a classical conventionalist. It is reasonable to interpret the conventionalist as endorsing this form of explanation.

world. So the form of the explanation would be the following: □p because @q (Necessarily p, because actually q).[227]

If this is the correct form of explanation for a contingentist to choose, then it doesn't follow any more that if q had not been the case, then p would not have been the case, because it is still true that actually q, which suffices to explain, according to the contingentist, that p is necessary. This way of reconstructing the form of explanation also fits the spirit of the contingentist proposal much better, for the contingentist is likely to also be an actualist, in the sense of believing that the facts of necessity (and possibility) are constrained by the way the actual world is.[228]

This form of explanation blocks an argument to the conclusion that the contingentist falls prey to Blackburn's contingency-horn. Lets see in some more detail how the argument, modeled on the anti-contingentist argument above, is blocked. We start with the same assumptions as above, with the form of explanation amended in premise (2*), and the counterfactual principle in (3*) amended accordingly.

(1) □p	Explanandum
(2*) □p because @q	Explanation of □p by q
(3*) ¬@q □→ ¬□p	Assumption: Plausible counterfactual principle about explanation
(4) ◊¬q	Assumption: Contingency of q

[227] Morato (2014) makes a similar proposal, by allowing for trans-world relations of explanation. However, the present solution is much simpler than the solution presented by Morato.

[228] We must distinguish between existence-actualism, and the actualism currently at issue, which I will call explanation-actualism. Existence-actualism is the thesis that only what is actual exists. The explanation-actualist is not committed to this strong thesis, for she can allow for the existence of possibilia, as long as their possibility is somehow grounded in the actual world.

From these premises nothing interesting follows. All of them are assumed to be true, but we cannot use (3*) together with (4) to conclude that p is possibly not necessary, for while it is possible that not q, it is not possible that not actually q. (3*) is still true on standard accounts for the counterfactual conditional, because counterfactuals with impossible antecedents are trivially true. Consequently, there is no valid argument to the conclusion that no explanation of necessity in terms of something contingent is possible, if the form of explanation the contingentist utilizes, or should utilize, is as outlined above.

This reconstruction of the argument also shows that the modal status of the explanans, q, does not play a role in the explanation of the necessity of p at all. Therefore, both the necessitist as well as the contingentist are off the hook. As long as q is actually true and the relevant explanatory relation holds, the explanation will not fall prey to the dilemma. Therefore, Blackburn's dilemma cannot be used to argue against positions that assume an explanans that is necessarily true, as long as the explanation does not depend on the modal status itself, which would most likely make the explanation unsatisfactory on independent grounds.[229]

I conclude that an argument against an explanation of necessary truth utilizing Blackburn's dilemma cannot get off the ground, if the correct form of explanation is used. The contingentist is not committed to rejecting any principles of modal logic, such as S4 and S5, and can equally accept a counterfactual principle about

[229] The moral, that the necessity of the explanans does nothing to show that necessity enters in an objectionable way into the explanation of itself, is also drawn by Hale (2002), who calls this kind of explanation 'non-transmissive-explanation'

explanation, for the necessitist, on the other hand, no bad 'residual must' remains.

The form of explanation in the current framework

The present theory claims that necessary truth is ultimately explained by relations holding between the requirements for truth of a representation and the reference realizing properties determined by the reference determining properties of a representation. To show that it is not vulnerable to a Blackburn-style argument, it needs to be shown that it does not exhibit an objectionable form of explanation.

What explains the necessary truth of a representations is the relation that holds between properties of the truth-evaluable representation and properties determined by the reference determining properties of the representation. This is not itself a representation or proposition, but just a relation between properties. To state the explanation, however, one needs to represent this fact and this representation can be necessarily or contingently true. So the question about the modal status of the explanans is whether a representation of the fact that a certain relation obtains between the reference realizing properties and the requirements for truth is itself necessarily true. So provided the explanans is a representation stating some relationship between properties of the truth-evaluable representation and properties of the referential representations, is it necessarily true? At first it may seem that it is contingent, because it is a contingent matter whether the relationship between these properties holds. After all, reference is a contingent matter.

But it is not quite as straightforward. For representations have their representational properties essentially and so a referential

representation cannot fail to determine the reference realizing properties and truth-evaluable representations cannot fail to have their requirements for truth. It remains a contingent matter to which objects a referential representation refers and it also remains a contingent matter what reference determining properties a certain word or thought has. But since representations are words cum meaning, or thoughts cum content, they have their representational properties essentially and so the relationship that holds between the requirements for truth and the reference realizing properties of the constituent referential representations holds necessarily. So contrary to what one may have expected, the explanans is necessarily true as well. Still, the explanation does not leave a residual must which would be in need of further explanation, because it is not the necessity of the explanans, but merely its truth that explains the necessity of the explanandum, and so necessity is not essentially involved in the explanation. Therefore the present explanation cannot be attacked using Blackburn's dilemma.

6. Explanation of Essence

A further reason for believing that the theory is a good explanation of necessary truth is that it allows for an account of what essences are. I already pointed out above that the account of essence provides some reason for taking the theory to be extensionally adequate, but the fact that it explains what essences are independently from its explanation of necessity gives rise to a further nice feature of the theory. For in this way, it is able to properly account for the relationship between essence and modality.

As Kit Fine famously pointed out, essence and necessity are closely connected, but it is not the case that essence can be explained by necessity.[230] Essence is a more fine-grained notion than necessity. Fine concluded from this that rather than explaining essence in terms of necessity, we should explain necessity in terms of essence.

The present explanation, in a sense, takes a third way, it accounts for what essences are and explains what necessity is in terms of reference realizing properties and requirements for truth. Essential properties just are the reference realizing properties determined by singular representations. This allows for both, an explanation of necessity and an account of essences that does not depend on the explanation of necessary truth.

Consequently, Fine's observation, which has convinced many of the merits of essentialism, can be taken seriously without having to take truth in virtue of essence as a primitive in the explanation of necessity.[231]

I take this to be a significant advantage over theories of necessity which do not have an independent way of accounting for essential properties, but have to rely on a modalist account of essences instead.

7. Minimal Assumptions

The final reason why the theory gives a good explanation of necessary truth is that it proceeds from minimal assumptions about representations generally. These assumptions should be fairly uncontroversial, if one accepts that there are representations

---

[230] Fine (1994)

[231] This is desirable, because primitive essences have trouble giving an illuminating explanation of how they give rise to necessary truth.

and that these representations are either referential or truth-evaluable.[232] To make this explicit, let me rehearse the assumptions from which the account proceeds.

I argued that when thinking about necessary truth, one should not merely think about language as one medium of representation, but about representations quite generally. So sentences and words are of interest to the theory only in so far as we conceive of them as representations: Objects which have their representational properties essentially.

I argued at length that in order to make the discussion independent of the idiosyncrasies of language, one has to move away from the communication model underlying much study of language and instead move to representations viewed generally. I will not rehearse this discussion here, but only state the main assumptions I made to characterize representations.

First we need to distinguish between referential representations and truth-evaluable representations. Truth-evaluable representations consist of referential representations and are true in virtue of the appropriate relationships between the referents of the constituent representations.

This assumption should not be very controversial, for it is closely related to a picture underlying the standard Fregean view of predicates as functions from extensions to truth-values. One may then view the reference of a predicate as those objects which deliver the truth value true. However, Frege's view, in a sense, skips the explanatory step of what gives rise to truth. Instead he opts for truth as an indicator of reference, and does not give an

---

[232] A radical skeptic about representation might disagree, and presently I have nothing to say to such a skeptic, safe that I take the skeptical position to be implausible on a global scale.

independent account of how reference is determined.[233] This is not objectionable, if the goal is to give an analysis of language and how the elements of sentences systematically work together, given that the sentences formed with a predicate express representations with certain truth values. But it should be clear that the direction of determination, on the Fregean picture, is from truth to reference, and not the other way around.[234]

If we care about how the truth of a truth-evaluable representation comes about and how this depends on the reference of the constituents, however, we cannot make reference depend on truth, rather we have to account for reference and then explain how this reference gives rise to truth.

Following this line of thought and the direction of explanation, an account of reference of referential representations independently from its function in giving rise to truth must be given. So here are the relevant, very general assumptions about reference: Referential representations refer to objects. Singular representations, if they refer at all, essentially refer to only one object, general representations may refer to more than one object. The holding of a reference relation between a representation and some object or objects depends on both, properties of the

---

[233] The doctrine that apart from reference there is a sense associated with a word which determines reference is an add on to the functional analysis without which it would still stand.

[234] I am not arguing that Frege, and philosophers following Frege, missed that point. Proponents of truth-conditional semantics, for example Davidson (1967), explicitly take truth to be prior to reference. Davidson's concern, however, is largely epistemic: how can we interpret other speakers' sentences given that they use them to speak the truth. So it may well be that the epistemic direction of determination (how we come to know which representation a sentence expresses) is from truth to reference, while the metaphysical direction of determination (how the truth of a representation depends on the reference of its constituents) runs from reference to truth.

representation, and properties of the referent. Both, the representation of an object and the object must work together to give rise to the reference relation. I called the properties of a representation which play this role its reference determining properties and the properties of the object which play this role the object's reference realizing properties. These properties are related, for the reference determining properties of a representation determine what the reference realizing properties of the object are, if a reference relation is established. This is not to say that the representation determines that the object which is the referent of the representation has these properties, the object has these properties quite independently from being referred to, it is just that if a reference relation holds between the representation and the referent, then the object has these properties.

With this concepts of reference realizing properties and requirements for truth in hand it is a small step to an explanation of necessary truth. For it may now happen that the reference realizing properties determined by the referential representations are related such that the requirements for truth just cannot fail to be satisfied: the referents must exhibit the right kind of relationship required for the truth in virtue of how the reference of the referential representations is determined.

The assumptions necessary for giving this explanation just follow from these basic considerations about reference and how reference and truth are related. The assumptions are quite minimal, because nothing specific has to be assumed about how reference really is determined for a specific referential representation or kind of representation. So in principle, the theory is compatible with any more specific theory of reference determination.

To sum up, the only assumptions needed for producing an explanation of necessary truth are first, that representations are the appropriate objects for which necessary truth should be explained, second that reference and truth are related such that truth (metaphysically) depends on reference, third, that the holding of the reference relation depends on properties of the representation and properties of the referent. With these ingredients an account of necessary truth follows quite naturally.

This concludes the discussion of why I take the present explanation of necessary truth to be a good philosophical explanation of necessary truth: It is a good explanation, because it is a non-circular, illuminating and intelligible, as well as general explanation of metaphysical necessity that appears extensionally adequate, accounts for puzzles about possibilia and necessary existence, explains what essences are, and, along the way, makes only minimal assumptions about the nature of representation and its connection to the world.

Conclusion

I have argued for an explanation of metaphysical necessity in terms of how reference determination and requirements for truth interact. I hope to have convinced you that the explanation is interesting and offers a range of explanatory benefits. I already used the last chapter to summarize the findings and to argue that the explanation is philosophically fruitful and illuminating. I want to use these concluding remarks to clarify the present theory's connection to analyticity, before giving an outlook of how the resources developed here may be put to use in answering questions related to the present investigation.

Analyticity and the present theory of necessity

The present explanation diverges quite radically from attempts to explain necessity in terms of analyticity in the twentieth century, but it has strong connections to an older account of analyticity: The Kantian account of analyticity in terms of concept-containment. This account served as a blueprint for the general explanation of necessary truth, because it revealed how truth may in certain cases be guaranteed by how the reference of the constituent representations is determined. The Kantian containment-account of analyticity, however, only describes a special case and as such is too narrow to account for necessary truth generally, and so I expanded and generalized the picture to be in principle applicable to all representations. It is in virtue of this ancestry I take the present theory of necessary truth to at the same time give an account of analytic truth: Analytic truth, in the wide sense, just is necessary truth.

The move back to a broadly Kantian picture of analyticity as applying not to sentences, but to representations, also reveals what has gone wrong with accounts of analyticity that took analyticity to be truth of a sentence in virtue of its meaning. Accounts of truth in virtue of meaning took the wrong objects to be analytic: sentences individuated independently of the representations they express. Thus, a confusion arose between sentences that have some element of broadly linguistic meaning which guarantees that the sentence type always express some true representation and sentences which express a representation that is guaranteed to be true. Proponents of analyticity fallaciously identified conventional truth with necessary truth, but a convention that a certain sentence-type should always be regarded as expressing some true representation just does not suffice for necessary truth.

That conventional truth and necessary truth do come apart is especially obvious in the case of sentences involving indexicals. Take the example "I am here". Given that the constituent indexical expressions function such that 'I' always denotes the speaker of a context, and 'here' always denotes the place of the context, the sentence cannot fail to express a true representation, because whenever it is uttered by any speaker it expresses a truth.[235] The truth it expresses, however, is not necessarily true. At each occasion of utterance, the sentence expresses a different representation which is contingently true. So legislating (directly or indirectly) that a sentence is always to express a true representation does not guarantee that it expresses a necessary

[235] The indexical expressions 'I' and 'here' almost certainly do not function according to this simple model in natural language, but for present purposes let us assume that it is true.

truth, for one may equally well adjust the interpretation of the sentence such that it expresses some contingently true representation on each occasion of use. This explains why analyticity conceived of as a property of sentences, or as truth in virtue of meaning, is not a useful concept, if the goal is to explain necessary truth.

We should thus distinguish between two different notions of analyticity, both of which have some historical precedent, but which have not always been clearly separated: Analyticity as a property of sentences and analyticity as a property of representations. The former notion of analyticity was prominent during much of the twentieth century, which is natural given the preoccupation with language in analytic philosophy during that time as well as the epistemic role analyticity was supposed to play.

The latter notion goes back to Kant, for whom analyticity is a property that applies to judgments, which should not be conceived of as sentences, but rather as representations.[236] Kant, of course, only had narrow analyticity in mind, analyticity as concept containment, but appropriately extending and generalizing the notion in an updated framework, leads to a wide notion of analyticity, which can be identified with necessary truth.

The epistemology of modality

There are two related questions that I have deliberately left open, because an adequate discussion would not have been possible here. The first is the epistemology of necessity, the second is

---

[236] I do not want to claim that Kant had the same notion of representation in mind as has been presented here. Nonetheless judgments are more similar to representations, than to mere sentences.

which specific (kinds of) representations turn out to be metaphysically necessary. The omission of the latter is a consequence of the former and so I will in this section be concerned with the epistemology of modality.

The epistemology of necessity, or of modality more generally, is a large field and to a certain extent it is independent of the metaphysics of modality. However, as Christopher Peacocke has argued quite convincingly, an epistemology and a metaphysics for a certain domain should meet what he calls "the integration challenge".[237] The challenge is to have an account of the epistemology of a domain which is fit to discover what the metaphysics say the phenomenon consists in. For if the epistemology and the metaphysics don't integrate, there is the danger that either our epistemic practices are not geared towards discovering what really is the case, and therefore may not be reliable (or only reliable by accident) or that our metaphysical account does not describe the phenomenon we intended to describe and which we are trying to discover with our epistemic practices.

In the special case of modality, our epistemic practices should therefore match the metaphysical account of what it takes for a truth to be necessarily true. Our practices for coming to know about what is (metaphysically) possible or necessary are rather diverse. We use thought experiments, we use real experiments, we use models, we use logic, we draw things, we imagine, we test for conceivability, and so on. All these methods, as is quite common in epistemology, are fallible for various reasons, but the search for metaphysical necessity and possibility is further complicated by

---

[237] Peacocke (1999)

the fact that there are different kinds of necessity and so we may, for example, sometimes confuse nomic necessity with metaphysical necessity, its stronger cousin, or we may confuse a still stronger kind of necessity, for example logical necessity with metaphysical necessity, or we may confuse any of these with epistemic necessity.

Traditionally, theories of necessity which appealed to analyticity had a seemingly persuading answer to the integration challenge: Since necessity is grounded in meaning, meaning is conventional, and conventions are known by the community of speakers, the necessary truths can be known by *a priori* reasoning from the conventions. That the story could not be quite so simple was shown by Kripke with his examples of *a posteriori* necessities, but it was still suggested by a few philosophers that the *a posteriori* character of these truths can be separated from the *a priori* principles which allegedly give rise to the necessity.[238] This seemingly straightforward answer to the integration challenge may be taken to speak in favor of the theory, not least because it accounts for the cogency of *a priori* reasoning and is probably a reason why some still hold on to modified versions of it. However, since the underlying theory of necessity in terms of conventional truth is not tenable, an epistemology based on it should not be readily accepted.

The present theory of modality does not lend this support to *a priori* reasoning, because the link between analyticity as a

---

[238] See for example Sidelle (1989), who explicitly argues that *a posteriori* necessities can be accounted for in this way by a conventionalist account. In a slightly different context, proponents of two-dimensional semantics like David Chalmers (2006) and Frank Jackson (2000) sometimes try to do something similar to defend the cogency of *a priori* reasoning in philosophy, although both do not endorse a conventionalist theory of necessity.

property of representations and *a priori* reasoning is severed, and so it does not share the epistemic payoff with its conventionalist predecessors.

But the present theory does leave us with some advice on how one may go about finding out about which representations are necessarily true. Since necessary truth is explained by the reference realizing properties of the constituent representations together with the requirements for truth, to know about the truth of a claim that a certain representation is necessarily true, we have to find out what the reference realizing properties determined by the referential representations are and what the requirements for truth are. This is something we may find out about in various ways. We already, for example, know quite a bit about the representations we use to communicate and about the representations we use to think about the world, so we may use this knowledge to assess necessity claims. But there is also room for a more sophisticated empirical study of representations that tells us something about how reference is determined and about what the reference realizing properties of some representations are. Also, linguistics may tell us something interesting about the representations expressed by the languages we use to communicate that can in turn be used to assess the necessity of a representation. Furthermore, scientific investigation may reveal what it really takes to be the member of a kind. And sometimes to come to know that a representation is necessarily true, it suffices to know something quite general about the representational system in which the representations occur: take for example the case of identities that are expressed by sentences involving the same name on either side of the identity sign. Since it is, with some exceptions, the case that the same name is used in the same

context to represent the same object, the two names are used as two representations of the same type, and so an identity expressed by the sentence is very likely to be necessarily true. In strongly regimented representational systems it may be even easier to come to know about necessarily true representations in this way.

So there are many ways in which one may come to know about the necessity of a representation and this first cursory glance at the possibilities of finding out about it seems to indicate that none of the standard methods for finding out about modality is excluded by the present theory. But some guidance is offered. The methods must be fit to reveal the features relevant to explaining necessity. In order to find out about necessity, one must investigate the reference realizing properties determined by the representations, the requirements for truth, and something about their relationship.

Thus there is some *prima facie* reason to believe that the epistemology the present theory of metaphysical necessity yields, integrates well with our *de facto* methods of finding out about modality, even if it does not deliver any validation or explanation of *a priori* reasoning.

Much more would have to be done to develop these sketchy observations, but for now I hope it gives a first glimpse as to how the present theory may integrate with the epistemology of modality and how it may be used to decide some conflicts about which representations are metaphysically necessary.

Essence and explanation

I argued that the present theory, apart from explaining necessary truth, can be used to elucidate what essences are. With this account of essence in hand, one may be able to go one step further

and get a better understanding of what philosophical or constitutive explanations are.

Recently the notion of grounding has gained some prominence in the philosophical debate. It is supposed to describe the relation holding between a fact and another which constitutively explains why the first obtains.[239] Above, when considering philosophical explanations, I did not directly use the notion of grounding, but instead appealed to a pre-theoretical distinction between constitutive and genealogical explanations. In a constitutive explanation we want to be told what it takes for something to be the case, or what it takes for something to exist. But what is this constitutive "what it takes to be for something to exist or for something to be the case" exactly? It seems that the account of essence above gives us some guidance at least for one of the two questions. For the essence of something tells us what it takes for something to exist. And so an explanation for why something (constitutively) exists, should cite the essential properties of the object in question, that is, it should cite the reference realizing properties determined by the singular referential representation of the object.

Given that the present account of essence in terms of reference realizing properties determined by singular representation is viable, the demands on a constitutive explanation more generally can be elucidated. For we can use the reference realizing properties of general referential representations to explain what it takes for something to be the case: What does it take, constitutively, to be a bachelor? There must be someone who is

---

[239] There is some disagreement whether grounding should be conceived of as a relation between facts. This debate can be neglected here. For more on this debate and on grounding generally see Correia and Schnieder (2012)

both male and unmarried. Quite generally we can give an answer to what it takes for something to be the case, by citing the reference realizing properties determined by the general referential representations used to represent a fact. This unifies the two constitutive explanations of what it takes for something to exist, and of what it takes for something to be the case: in both cases an answer can be given by citing the reference realizing properties determined by the representations in question.

If this connection between reference realizing properties and constitutive explanation can be sustained it would provide a unified understanding of essence, necessity, and constitutive philosophical explanation, which would be quite desirable.

An example where this unity can be put to use is a debate in the discussion on grounding. There a debate has emerged about whether grounding can be explained in terms of essence, and counterexamples seem to suggest that this is not the case.[240] The present understanding of philosophical explanation can agree that essence does not explain grounding generally, for it may not do so when we ask what it takes for something to be the case. However, one can insist that philosophical explanations should generally cite reference realizing properties. If we want an explanation of why something exists, we cite the reference realizing properties determined by the singular representation, that is, the essence of the object. If we want an explanation of why something is the case, we cite the reference realizing properties determined by the general representation. So a common feature, reference realizing properties, explains why essence and grounding are related, but essence fails to explain every grounding explanation.

---

[240] For a defense of an analysis of grounding in terms of essence see Correia (2013), for a dissenting view see Fine (2012)

This unified picture of necessity, essence, and constitutive explanation may thus prove useful for further research, but more will have to be done to show that the initially appealing unity can be upheld.

Metametaphysics - internal and external questions

Rudolf Carnap made an influential distinction between internal existence questions and external existence questions.[241] The internal existence questions are of the kind: "Is there beer in the fridge", or "Does the Higgs Boson exist?". They can be answered by straightforward empirical investigation, even if this empirical investigation may be quite complicated and the questions have a clear positive or a negative answer. These internal questions are, according to Carnap, to be distinguished from external questions, which cannot be answered in this way, but are to be answered by pragmatic considerations, considerations about whether talk about the type of entity introduced is useful to the (scientific) enterprise in question. Many philosophical existence questions, for example "Are there numbers?" or "Do abstract entities exist?", fall in this category according to Carnap, and should be answered accordingly by considering whether the adoption of a linguistic framework in which we quantify over these entities is pragmatically acceptable. Carnap introduced the distinction with the explicit intention to make talk of abstract entities acceptable to empiricists and today the distinction is equally put to use to deflate metaphysical debates about the existence of, for example, composite objects.

---

[241] Carnap (1950)

However, the distinction was criticized right after its introduction, and modern versions of it are criticized for similar reasons. Quine, for example, notes that it seems that external questions can simply be turned into internal questions by adopting a linguistic framework which allows us to quantify over the entities in question. So once we have a language that quantifies over numbers, the question "Are there numbers?" becomes an internal question and no longer a pragmatic choice over what we should talk about.[242] And so why could we not have a linguistic framework which covers everything there is, and consider this to be the framework in which we ask philosophical questions of ontology? After all, philosophers doing ontology typically ask what there is in the widest possible sense, they do not just ask what is in the fridge.

This criticism is rightly taken to be quite serious for the advertised deflationary understanding of ontological questions: Unwanted entities cannot just be excluded from one's ontology by choosing a linguistic framework such that it excludes talk about some unwanted entities. However, something intuitively right about the distinction can possibly be salvaged with the help of the present theory. For the tools of the theory make sense of the intuition behind the distinction that before we can meaningfully ask whether something exists, we should have an idea of what it takes for this something to exist. What it takes for something to exist is that it has the reference realizing properties determined by the representation representing it. This makes sense of the internal question. Once we know what it takes for something to exist, we

[242] See Quine (1951b) p. 69

can ask whether something has what it takes to be the thing in question.

The external questions can then be conceived of as truly pragmatic questions about what of the many things in the world should be represented for the specific purposes at hand. With this understanding, however, philosophical questions about whether numbers exist or whether there are abstract objects, will not fall into the purview of external questions. Rather, what is required to answer them is a clarification of what it would take for numbers to exist or what it would take for abstract objects to exist and to see whether what it takes is indeed the case.

However, some existence questions could still be read as being truly external. If a biologist, for example, asks whether there really are species, then the question may be construed as a question about whether it is useful for the biological enterprise to distinguish between animals along the line of species, or whether it might not be more fruitful to draw a different distinction. Equally, the sociologist might ask whether classes exist, because she believes that an analysis of societies along different dimensions is more fruitful. In both cases, it would be quite wrong to interpret the biologist as literally saying that there is no such thing as, say, dogs, or the sociologist as saying, that there is no working-class, rather, what they are saying is that neither species nor classes are very helpful in explaining what we want to explain in the relevant discipline, and that we should therefore look at different entities.

This understanding of the internal-external distinction flows quite naturally from the understanding of representations and the reference realizing properties: The reference realizing properties are properties of an object and whether some object has these

properties is not a matter of whether they are represented by anything. So whether something exists is a matter that must be investigated by seeing whether anything really has what it takes for the object in question to exist. But that these particular properties are the properties constitutive of its existence, that they are what it takes for the object to exist, is determined by the representation used to talk about that object. In this sense, external existence questions depend on what we choose to represent. Talk of choice may be somewhat of an overstatement, however, for many times it is not really a matter of choice for us what the best way of talking about something is. This may rather be dictated by the way the world really is or by what a community of speakers in fact represents. Nonetheless, in the sense brought out here, it is a pragmatic question what we choose to represent.

Making sense of the distinction in this way may bring some light into some philosophical debates about the existence of, say, abstract objects, for it may be that proponents on either side of the debate do not so much disagree about the internal question of whether anything has the reference realizing properties determined by the representation in question, they disagree about what it would take for abstract objects to exist. One side may take a quite undemanding view, another may require so much that it rightly seems impossible that it is satisfied by anything. By seeing what it is that is really represented by the representations allegedly referring to abstract objects, we can seek to resolve such disputes.

The end

The remarks in this conclusion show where the present explanation of necessary truth, and the foundation on which it

rests leaves room for further investigation, but I think it also shows that the move away from sentences and propositions in the traditional language-inspired framework, to a framework of representations and their relation to what they represent, apart from explaining necessary truth, has some promise to give fruitful answers to intricate philosophical questions: It may allow for advances in the epistemology of modality, it could potentially provide a unified picture of necessity, essence, and ground, and it gives some basis to the Carnapian intuition of internal and external questions, thus allowing to make better sense of age-old philosophical conundrums.

I argued in some detail that one of the promises of the framework of representations, reference, and truth, holds true; the promise that it explains necessary truth. As for the others, I must now close the book, leaving the issues open.

Bibliography

Adams, Robert M., (1974), "Theories of Actuality", in *Noûs*, Vol. 8, No. 3. p. 211-231.

Almotahari, Mahrad, (2014), "The Identity of a Material Thing and its Matter", in *The Philosophical Quarterly*, Vol. 64, No. 256, p. 387-406.

Anderson, Lanier R., (2015), "The Poverty of Conceptual Truth, Kant's Analytic/Synthetic Distinction and the Limits of Metaphysics", Oxford: Oxford University Press.

Armstrong, D.M., (1983), "What is a Law of Nature?", Cambridge: Cambridge University Press.

Armstrong, D.M., (1993), "The Identification Problem and the Inference Problem", in *Philosophy and Phenomenological Research*, Vol. 53, No. 2, p. 421-422.

Ayer, Alfred J., (1936a), "Truth by Convention: A Symposium by A.J. Ayer, C.H. Whiteley, M. Black", in *Analysis*, Vol. 4, No. 2/3, p. 17-32.

Ayer, Alfred J., (1936b), "Language, Truth and Logic", New York: Oxford University Press.

Balaguer, Mark, (2015), "Fictionalism in the Philosophy of Mathematics", in *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://

plato.stanford.edu/archives/sum2015/entries/fictionalism-mathematics/>.

Beaney, Michael, (2014) "Analysis", in *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/sum2016/entries/analysis/>

Blackburn, Simon, (1993), "Morals and Modals" in *Essays in Quasi-Realism*, Simon Blackburn ed., Oxford: Oxford University Press.

Boghossian, Paul A., (1996), "Analyticity Reconsidered", in *Nôus*, Vol. 30, No. 3, p. 360-391.

Borghini, Andrea; Williams, Neil E., (2008), "A Dispositional Theory of Possibility", in *Dialectica*, Vol 62, No. 1, p. 21-41.

Bromberger, Sylvain, (2011), "What are words? Comments on Kaplan (1990), on Hawthorne and Lepore, and on the Issue", in *The Journal of Philosophy*, Vol. 108, No. 9, p. 486-503.

Burge, Tyler, (1979a), "Sinning Against Frege", in *The Philosophical Review*, Vol. 88, No. 3, p. 398-432.

Burge Tyler, (1979b), "Individualism and the Mental", in *Midwest Studies in Philosophy*, Vol. 4, Issue 1, p. 73-121.

Burgess, Alexis; Sherman, Brett, (2014), "Metasemantics: New Essays on the Foundations of Meaning", Oxford: Oxford University Press.

Cameron, Ross, (2010a), "On the Source of Necessity", in *Modality: Metaphysics, Logic, and Epistemology*, Bob Hale and Aviv Hoffman (eds.), Oxford: Oxford University Press.

Cameron, Ross, (2010b), "The Grounds of Necessity", in *Philosophy Compass*, Vol. 5, No.4, p. 348-358.

Cameron, Ross, (2012), "Why Lewis's Analysis of Modality Succeeds in its Reductive Ambitions", in *Philosopher's Imprint*, Vol. 12, No. 8, p. 1-21.

Carnap, Rudolf, (1952), "Meaning Postulates", in *Philosophical Studies*, Vol. 3, No. 5, p. 65-73.

Carnap, Rudolf, (1947), "Meaning and Necessity", Chicago: University of Chicago Press. Page References to the Midway reprint of the second edition (1988).

Carnap, Rudolf, (1950), "Empiricism, Semantics, and Ontology", in *Revue Internationale de Philosophie*, Vol. 4, No. 11. p. 20-40.

Chalmers, David J., (2006), "Two-Dimensional Semantics", in *Oxford Handbook of the Philosophy of Language,* Ernie Lepore and Barry Smith (eds.). Oxford: Oxford University Press, p. 574-606.

Chalmers, David, J., (2012), "Constructing the World", Oxford: Oxford University Press.

Cohen, Jonathan; Michaelson Eliot, (2013), "Indexicality and The Answering Machine Paradox", in *Philosophy Compass*, Vol.8, Issue 6, p. 580-592.

Correia, Fabrice, (2006), "Generic Essence, Objectual Essence, and Modality" in *Noûs*, Vol. 40, No. 4, p. 753-767.

Correia, Fabrice; Schnieder Benjamin (eds.), (2012), "Metaphysical Grounding: Understanding the Structure of Reality", Cambridge: Cambridge University Press.

Correia, Fabrice, (2012), "On the Reduction of Necessity to Essence", in *Philosophy and Phenomenological Research*, Vol. 84, No. 3, p. 639-653.

Correia, Fabrice, (2013), "Metaphysical Grounds and Essence", in *Varieties of Dependence. Ontological Dependence, Grounding, Supervenience, Response-Dependence.* Hoeltje, M.; Schnieder, B; Steinberg A. (eds.), Basic Philosophical Concepts Series, München: Philosophia, p. 271-296.

Davidson, Donald, (1967), "Truth and Meaning", in *Synthese*, Vol. 17, p. 304-324.

Devitt, Michael; Sterelny Kim, (1999), "Language and Reality, An Introduction to the Philosophy of Language", 2nd Edition, Cambridge Mass.: The MIT-Press.

Divers, John; Melia, Joseph, (2002), “The Analytical Limit of Genuine Modal Realism”, in *Mind*, Vol. 111, No. 441, p. 15-36.

Divers, John; Melia, Joseph, (2003), “Genuine Modal Realism Limited”, in *Mind*, Vol. 112, No. 445, p. 83-86.

Divers, John; Melia, Joseph, (2006), “Genuine Modal Realism: Still Limited”, in *Mind*, Vol. 115, No. 459, p. 731-740.

Divers, John, (2014), “The Modal Status of the Lewisian Analysis of Modality”, in *Mind*, Vol. 123, No. 491, p. 861-872.

Evans, Gareth, (1973), “The Causal Theory of Names”, in P*roceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 47. p. 187-225.

Evans, Gareth, (1982), “The Varieties of Reference”, Oxford: Clarendon Press.

Field, Hartry, (1989), “Realism, Mathematics and Modality”, Oxford: Basil Blackwell.

Fine, Kit, (1994), “Essence and Modality”, in *Philosophical Perspectives*, Vol. 8, Logic and Language, p. 1-16.

Fine, Kit, (1995), “Senses of Essence”, in *Modality, Morality, and Belief, Essays in Honor of Ruth Barcan Marcus*, Sinnott-Armstrong, Walter (ed.), Cambridge: Cambridge University Press, p. 53-73.

Fine, Kit, (2001), “The question of realism”, in *Philosopher’s Imprint*, Vol.1, No.1, p. 1-30.

Fine, Kit, (2002), “Varieties of Necessity”, in *Conceivability and Possibility*, Gendler, Tamar Szabo; Hawthorne, John (eds.), Oxford: Oxford University Press, p. 253-281.

Fine, Kit, (2003), “The Non-Identity of a Material Thing and its Matter”, in *Mind*, Vol. 112, No.446, p. 195-234.

Fine, Kit, (2005), “Modality and Tense: Philosophical Papers”, Oxford: Oxford University Press.

Fine, Kit, (2012), “Guide to Ground”, in *Metaphysical Grounding*, Correia, Fabrice; Schnieder, Benjamin (eds.), Cambridge: Cambridge University Press, p. 37-80.

Fine, Kit, (2015), “Unified Foundations for Essence and Ground”, in *Journal of the American Philosophical Association*, Vol. 1, No. 2, p. 296-311.

Fine, Kit, (2016), “Identity Criteria and Ground”, in *Philosophical Studies*, 173, p. 1-19.

Frege, Gottlob, (1884), “Die Grundlagen der Arithmetik. Eine logisch mathematische Untersuchung über den Begriff der Zahl”, Breslau: Verlag von Wilhelm Koebner. Page references according to: (1988), Christian Thiel (ed.), Hamburg: Meiner Verlag.

Frege, Gottlob, (1948), "Sense and Reference", in *The Philosophical Review*, Vol.57, No. 3, Max Black (transl.), p. 209-230.

García-Carpintero, Manuel; Pérez Otero, Manuel, (2009), "The Conventional and the Analytic", in *Philosophy and Phenomenological Research*, Vol. 78, No. 2, p. 239-274.

Hale, Bob; Wright, Crispin, (1994), "A Reductio Ad Surdum? Field on the Contingency of Mathematical Objects", in *Mind*, Vol 103, No. 410, p. 169-184.

Hale, Bob, (2002), "The Source of Necessity", in *Philosophical Perspectives*, Vol.16, p. 299-319.

Hale, Bob; Hoffmann, Aviv (eds.), (2010), "Modality: Metaphysics, Logic, and Epistemology", Oxford: Oxford University Press.

Hale, Bob, (2013), "Necessary Beings: An Essay on Ontology, Modality, and the Relations Between Them", Oxford: Oxford University Press.

Hanks, Peter, (2008), "A Dilemma About Necessity", in *Erkenntnis*, Vol. 68, p. 129-148.

Hanks, Peter, (2009), "Recent Work on Propositions", in *Philosophy Compass*, Vol. 4, No. 3, p. 469-486.

Harman, Gilbert, (1996), “Analyticity Regained?”, in *Nôus*, Vol. 30, No. 2, p. 392 - 400.

Hawthorne, John; Lepore, Ernest, (2011), “On Words”, in *The Journal of Philosophy*, Vol. 108, No.9, p. 447-485.

Heim, Irene; Kratzer, Angelika, (1998), “Semantics in Generative Grammar”, Oxford: Blackwell Publishing.

Hellman, Geoffrey, (1989), “Mathematics without Numbers”, Oxford: Oxford University Press.

Hofmann, Frank; Horvath Joachim, (2008), “In Defence of Metaphysical Analyticity”, in *Ratio*, Vol. 21, p. 300-313.

Jackson, Frank, (2000), “From Metaphysics to Ethics. A Defense of Conceptual Analysis”, Oxford: Oxford University Press.

Jubien, Michael, (2009), “Possibility”, Oxford: Oxford University Press.

Kant, Immanuel, (1998), “Critique of Pure Reason”, Paul Guyer and Allen Wood (eds. and translation), Cambridge: Cambridge University Press.

Kaplan, David, (1977), “Demonstratives”, in *Themes from Kaplan*. Joseph Almog, John Perry, and Howard Wettstein (eds.), Oxford: Oxford University Press. p. 481-563.

Kaplan, David, (1990), "Words", in *Proceedings of the Aristotelian Society Supplementary Volume*, Vol. 64, p. 93-119.

Kennedy, Christopher, (2006), "Comparatives, Semantics of", in *Encyclopedia of Language and Linguistics*, second Edition, Brown, Keith (ed.), Oxford: Elsevier, p. 690-694.

Keefe, Rosanna, (2002), "When does Circularity Matter?" in *Proceedings of the Aristotelian Society*, Vol. 102, Issue 1, p. 275-292.

King, Jeffrey C.; Soames, Scott; Speaks, Jeff., (2014), "New Thinking about Propositions" Oxford: Oxford University Press.

Klein, Ewan, (1980), "A Semantics for Positive and Comparative Adjectives", in *Linguistics and Philosophy*, Vol. 4. p. 1-45.

Koslicki, Kathrin, (2012), "Essence, necessity, and explanation", in *Contemporary Aristotelian Metaphysics*, Tahko, Tuomas (ed.), Cambridge: Cambridge University Press, p. 187-206.

Kripke, Saul, (1959), "A Completeness Theorem in Modal Logic", in *Journal of Symbolic Logic*, No. 24, p. 1-14.

Kripke, Saul, (1963), "Semanitcal Considerations on Modal Logic", in *Acta Philosophica Fennica*, No. 16, p. 83-94.

Kripke, Saul, (1980), "Naming and Necessity", Cambridge Mass.: Harvard University Press.

Lange, Marc, (2008), “Why Contingent Facts cannot Necessities Make”, in *Analysis*, Vol. 68, No. 2, p. 120-128.

Lange, Marc, (2009), “Laws and Lawmakers”, Oxford: Oxford University Press.

Langford, C.H., (1942), “The Notion of Analysis in Moore’s Philosophy”, in *The Philosophy of G.E. Moore*, Schilpp, P.A. (ed.), p. 321-342.

Leech, Jessica, (2015), “The Varieties of (relative) Modality”, in *Pacific Philosophical Quarterly*, Vol. 97, Issue 2, p. 158-180.

Leech, Jessica, (forthcoming 2016), “The Mereology of Representations”, in *Proceedings of the Aristotelian Society*, Vol. 116, No. 2.

Lewis, David K., (1973), “Counterfactuals”, Victoria: Blackwell Publishing.

Lewis, David K., (1975), “Languages and Language”, in *Minnesota Studies in the Philsophy of Science*, Gundersond, Keith (ed.), Minneapolis: University of Minnesota Press. p. 3-35.

Lewis David K., (1983), “Philosophical Papers Vol I”, Oxford: Oxford University Press.

Lewis, David K., (1986), “On the Plurality of Worlds”, Victoria: Blackwell Publishing.

Lewis David K., (1987), “Philosophical Papers Vol.II”, Oxford: Oxford University Press.

Lewis, David K., (1996), “Elusive Knowledge”, in *Australasian Journal of Philosophy*, Vol.74, Issue 4, p. 549-567.

Lewy, Casimir, (1976), “Meaning and Modality”, Cambridge: Cambridge University Press.

Loewer, Barry M., (1997), “A Guide to Naturalizing Semantics”, in *A Companion to the Philosophy of Language*, Wright, Crisptin; Hale, Bob (eds.), Oxford: Blackwell Publishing, p. 108-126.

Lowe, E.J., (2001), “The Possibility of Metaphysics: Substance, Identity, and Time”, Oxford: Oxford University Press.

Lowe, E.J., (2008a), “Essentialism, Metaphysical Realism, and the Errors of Conceptualism”, in *Philosophia Scientiae*, Vol 12, No. 1. p. 9-33.

Lowe, E.J. (2008b), “Two Notions of Being: Entity and Essence”, in *Royal Institute of Philosophy Supplement*, 62, p. 23-48.

Mackie, Penelope, (2006), “How Things Might Have Been”, Oxford: Oxford University Press.

Marcus, Ruth Barcan, (1947), “Identity of Individuals in a Strict Functional Calculus of Second Order”, in *Journal of Symbolic Logic*, Vol. 12, p. 12-15.

Marcus, Ruth Barcan, (1995), “Modalities”, Oxford: Oxford University Press.

Menzel, Christopher, (2016) “Actualism”, in *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2016/entries/actualism/>.

Millikan, Ruth, (1989), “Biosemantics”, in *The Journal of Philosophy*, Vol. 86, p. 281-297.

Moltmann, Friederike, (2013), “Reference to Numbers in Natural Language”, in *Philosophical Studies*, Vol. 162, Issue 3, p. 499-536.

Moore, G.E., (1903), “Principia Ethica”, Cambridge: Cambridge University Press.

Moore, G.E., (1942), “A Reply to my Critics”, in *The Philosophy of G.E. Moore*, Schilpp, P.A. (ed.), p. 660-667.

Morato, Vittorio, (2014), “Explanation and Modality: On the Contingency Horn of Blackburn’s Dilemma”, in *Erkenntnis*, Vol. 79, Issue 2, p. 327-349.

Olson, Eric T., (2016) “Personal Identity”, in *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/identity-personal/>

Paul, Laurie A., (2006), “In Defense of Essentialism”, in *Philosophical Perspectives*, Vol. 20, Issue 1, p. 333-372.

Peacocke, Christopher, (1999), “Being Known”, Oxford: Oxford University Press.

Plantinga, (1978), “The Nature of Necessity”, Oxford: Oxford University Press.

Putnam, Hilary, (1967), “Mathematics without Foundations”, in *The Journal of Philosophy*, Vol. 64, No.1, p. 5-22.

Putnam, Hilary, (1973), “Meaning and Reference”, in *The Journal of Philosophy*, Vol. 70, No. 19, p. 699-711.

Putnam, Hilary, (1975) “The Meaning of Meaning”, in *Minnesota Studies in the Philosophy of Science*, Vol. 7. p. 131-193.

Quine, Willard van Orman, (1951a), “Two Dogmas of Empiricism”, in *The Philosophical Review*, Vol. 60, No. 1. p. 20-43.

Quine, Willard van Orman, (1951b), “On Carnap’s Views on Ontology”, in *Philosophical Studies*, Vol. 2, No. 5, p.65-73.

Quine, Willard van Orman, (1960), “Carnap on Logical Truth”, in *Synthese*, Vol. 12, No.4, p. 350-374.

Quine, Willard van Orman, (1976), “Truth by Convention” in *The Ways of Paradox and other Essays*, Cambridge, Mass.: Harvard University Press.

Quine, Willard van Orman, (1986), “Philosophy of Logic”, Cambridge Mass.: Harvard University Press.

Rey, Georges, (2013) “The Analytic/Synthetic Distinction”, in *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2015/entries/analytic-synthetic/>

Roca-Royes, Sonia, (2011a), “Essentialism vis-à-vis Possibilia, Modal Logic and Necessitism” in *Philosophy Compass*, Vol. 6, Issue 1, p. 54-64.

Roca-Royes, Sonia, (2011b), “Essential Properties and Individual Essences”, in *Philosophy Compass*, Vol. 6, Issue 1, p. 65-77.

Rosen, Gideon, (2010), “Metaphysical Dependence: Grounding and Reduction”, in *Modality: Metaphysics, Logic, and Epistemology*, Hale, Bob; Hoffmann, Aviv (eds.), Oxford: Oxford University Press.

Russell, Bertrand, (1905), “On Denoting”, in *Mind*, Vol. 14, No. 56, p. 479-493.

Russell, Gillian, (2008), “Truth in Virtue of Meaning, A Defense of the Analytic/Synthetic Distinction”, Oxford: Oxford University Press.

Russell, Gillian, (2010), “A new Problem for the Linguistic Doctrine of Necessary Truth”, in *New Waves in Truth*, Wright, C., Pedersen, N. (Eds.), London: Palgrave Macmillan. p. 267-281.

Salmon, Nathan, (1981), “Reference and Essence”, Princeton: Princeton University Press

Salmon, Nathan, (1983), “Frege’s Puzzle”, Cambridge Mass.: The MIT-Press

Salmon, Nathan, (1993), “Analyticity and Apriority”, in *Philosophical Perspectives*, Vol. 7, p. 125-133

Schnieder, Benjamin, (2011), “A Logic for ‘Because’”, in *The Review of Symbolic Logic*, Vol. 4, No.3, p. 445-465.

Searle, John R., (1983), “Intentionality”, Cambridge: Cambridge University Press.

Shalkowski, Scott, (1994), “The Ontological Ground of the Alethic Modality”, in *The Philosophical Review*, Vol. 103, No. 4, p. 669-688.

Shoemaker, Sydney, (1998), “Causal and Metaphysical Necessity”, in *Pacific Philosophical Quarterly*, Vol. 79, p. 59-77.

Sidelle, Alan, (1989), “Necessity, Essence, and Individuation: A Defense of Conventionalism”, Ithaca NY: Cornell University Press.

Sidelle, Alan, (1991), "The Answering Machine Paradox", in *Canadian Journal of Philosophy*, Vol. 21, No. 4, p. 525-539.

Sidelle, Alan, (2009), "Conventionalism and the Contingency of Conventions", in *Noûs*, Vol. 43, No. 2, p. 224-241.

Sider, Ted, (2003), "Reductive Theories of Modality", in *The Oxford Handbook of Metaphysics*, M.J. Loux, D.W. Zimmermann, (eds.), Oxford: Oxford University Press, p. 180-208.

Stalnaker, Robert, (2003), "Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays", Oxford: Oxford University Press.

Stanford, Kyle; Kitcher Philip, (2000), "Refining the Causal Theory of Reference for Natural Kind Terms" in *Philosophical Studies*, Vol. 97, p. 99-129.

Strawson, P.F., (1950), "On Referring", in *Mind*, Vol. 59, No. 235, p. 320-344.

Tennant, Neil, (1997), "On the Necessary Existence of Numbers", in *Noûs*, Vol. 31, No. 3. p. 307-336.

Thomasson, Amie L., (2007), "Ordinary Objects", New York: Oxford University Press.

Thomasson, Amie L., (2013), “Norms and Necessity”, in *The Southern Journal of Philosophy*, Vol. 51, No. 2, p. 143-160.

Thomasson, Amie L., (2015), “Ontology made Easy”, New York: Oxford University Press.

Trogdon, Kelly, (2013), “An Introduction to Grounding”, in *Varieties of Dependence. Ontological Dependence, Grounding, Supervenience, Response-Dependence.* Hoeltje, M.; Schnieder, B; Steinberg A. (eds.), Basic Philosophical Concepts Series, München: Philosophia, p. 97-122.

Vetter, Barbara, (2015), “Potentiality: From Dispositions to Modality”, Oxford: Oxford University Press.

Wang, Jennifer, (2013a), “Modal Primitivism”, PhD Thesis, Published Online at Rutgers University Community Repository: http://dx.doi.org/doi:10.7282/T36T0K6Z

Wang, Jennifer, (2013b), “From Combinatorialism to Primitivism”, in *Australasian Journal of Philosophy*, Vol. 91, Issue 3, p. 535-554.

Wiggins, David, (2001), “Sameness and Substance Renewed”, Cambridge: Cambridge University Press.

Wiggins, David, (2012), “Identity, Individuation and Substance”, in *European Journal of Philosophy*, Vol. 20, No.1, p. 1-25.

Wildman, Nathan, (2013), “Modality, Sparsity, and Essence”, in *Philosophical Quarterly*, Vol. 63, No. 253, p.760-782.

Williamson, Timothy, (2002), “Knowledge and its Limits”, Oxford: Oxford University Press.

Williamson, Timothy, (2006), “Conceptual Truth”, in *Aristotelian Society Supplementary Volume 80*, No. 1, p. 1-41.

Williamson, Timothy, (2010), “Necessitism, Contingentism, and Plural Quantification”, in *Mind*, Vol. 119, p.657-748.

Williamson, Timothy, (2013), “Modal Logic as Metaphysics”, Oxford: Oxford University Press.

Wilson, Jessica, (2014), “No Work for a Theory of Grounding”, in *Inquiry: An interdisciplinary Journal of Philosophy,* Vol. 57, No. 5-6, p. 535-579.

Yablo, Stephen, (1992), “Review of: Necessity, Essence, and Individuation: A Defense of Conventionalism”, in *The Philosophical Review*, Vol. 101, No. 4, p. 878-881.